VOLUME XX[A] JULY, 1928 PARTS I AND II

BIOMETRIKA

ON THE NEED FOR STANDARDISATION IN MEASUREMENTS ON THE LIVING.

BY P. C. MAHALANOBIS, M.A. (Cantab.), Professor,
Presidency College, Calcutta.

I. INTRODUCTION.

1. LACK of agreement between different observers about fundamental definitions and technique of measurement constitutes an almost insurmountable obstacle to comparative studies in Anthropometry of the living. In an attempt to discover how far the measurements provided by different workers can be used for purposes of inter-racial comparisons I selected the following material:

(i) Y. Koganei (1893). "Beiträge zur physischen Anthropologie der Aino."

(ii) Aleš Hrdlička (1912). "The Natives of Kharga Oasis, Egypt."

(iii) F. von Luschan (1913). "Beiträge zur Anthropologie von Kreta."

(iv) T. Kubo (1913). "Beiträge zur phys. Anthropologie der Koreaner."

(v) Fritz Sarasin (1916—1922). "Anthropologie der Neu-Caledonier und Loyalty-Insulaner."

(vi) L. H. Dudley-Buxton (1922). "Ethnology of Malta and Gozo."

(vii) S. M. Shirokogoroff (1923). "Anthropology of Northern Asia."

(viii) S. M. Shirokogoroff (1925). "Anthropology of Eastern China, etc."

(ix) Aleš Hrdlička (1925). "Old Americans."

(x) H. Lundborg and F. J. Linders (1926). "Anthropologica Suecica."

2. All the above-mentioned authors are well-known workers in anthropology. Koganei was the Professor of Anatomy in the Imperial University of Tokio, and Kubo the Professor of Anatomy in Taihan Hospital, Seoul, Korea, at the time they took their measurements on the Aino (8)* and Koreans (10) respectively. Hrdlička was the Curator (Division of Physical Anthropology) at the United States National Museum, when he collected the materials for his monograph on the Egyptians of Kharga Oasis (4). F. von Luschan and Fritz Sarasin were trained as anthropologists and have published numerous important papers and books on anthropology. H. Lundborg and F. J. Linders are the Director and Vice-Director respectively of

* A list of bibliographical references is given in Appendix II. Numbers within brackets refer to that list.

the Swedish State Institute for Race Biology in Upsala. Dudley-Buxton was a member of the staff of the Department of Human Anatomy in the University of Oxford when he collected the material for his paper on the Maltese (3), and is at present, I believe, the Head of the Department of Anthropology in the same University. Shirokogoroff has published several papers on the comparative anthropology of the Far East, and appears to be an acknowledged authority in this subject.

3. One would naturally expect that the works of such trained scientists would be free from ambiguities, and would be comparable with one another, especially when almost all the investigations (with one exception, that of Koganei) were conducted several years later than the date of the International Agreement of Monaco on Anthropometric Measurements in 1906 (20). To my great surprise and consternation I found that this was far from the actual fact. Fresh discrepancies in definitions or in technique continually cropped up, and I was forced to make a systematic comparison of the different definitions of the measurements on the flesh. In the present paper I have given the results of such a comparison, restricting myself to measurements on the living head, and to the following standard lists cited in one or other of the publications mentioned above.

(i) "Notes and Queries on Anthropology" (1st edition, 1874).

(ii) Paul Broca (1879). "Instructions générales pour les recherches anthropologiques à faire sur les vivants."

(iii) Paul Topinard (1885). "Éléments d'Anthropologie générales."

(iv) Paul Topinard (English translation, 1890). "Anthropology."

(v) R. Virchow (1885). List on pp. 99—102 of "Zeitschrift für Ethnologie, Verhandlungen," Bd. XVII.

(vi) E. Schmidt (1888). "Anthropologische Methoden."

(vii) "Notes and Queries" (2nd edition, 1892).

(viii) "Schedule of Measurements in Anthropology" (British Association, 1895).

(ix) "Notes and Queries" (3rd edition, 1899).

(x) F. von Luschan (1906). Article on physical anthropometry in Neumayer's "Anleitung zu wissen. Beobachtungen auf Reisen," Bd. II.

(xi) Report of the Anthropometric Commission appointed by the XIIIth International Congress of Prehistoric Anthropology and Archaeology at Monaco, 1906.

(xii) Report of the Anthropometric Committee of the British Association (1908).

(xiii) "Notes and Queries" (4th edition, 1912).

(xiv) Rudolf Martin (1914). "Lehrbuch der Anthropologie."

(xv) Aleš Hrdlička (1920). "Anthropometry."

4. Koganei gives very short definitions, which are usually insufficient, and refers to Virchow's paper (31). Unfortunately this paper gives merely a list of measurements without any definitions and descriptive notes. Koganei however also mentions Schmidt's book (25), and it is probable that he actually followed it as his standard.

Kubo, in his work on the Koreans (10), constantly refers to Schmidt (25) and occasionally to von Luschan (11). Kubo was a pupil of Koganei's (Kubo mentions this fact in dedicating his book to Koganei), it therefore appears probable that he had been introduced to Schmidt's list by Koganei himself.

Hrdlička in his work on the Egyptians (4) does not give any systematic definitions, neither does he refer to any standard list. The system of measurements described in his book on Anthropometry (5) was published several years later, in 1920, but in this latter book he mentions that "the procedures, instruments, etc., to be here described, are those in regular use at the Division of Physical Anthropology, U.S.A. National Museum and in the field-work for the same" ((5), p. 33). As Hrdlička had gone out to Egypt while holding the post of Curator of the same department it may be presumed that his measurements were taken in accordance with the specifications described in his book on Anthropometry (5). The work on "Old Americans" (6), published in 1925, may also be presumed to be based on the same standards.

F. von Luschan, in his paper on the Cretans, distinctly mentions that "his technique was the same as that which he had originally learnt from Broca and which he himself, with slight modifications, has taught for three decades in practical courses" (translated from S. 354, *Zeitschrift für Ethnologie*, 1913). Broca's instructions (1) and von Luschan's own instructions in (11) have been consulted for elucidating his definitions.

Sarasin does not give any systematic definitions, but in view of his numerous references to Martin (14), I have assumed that Sarasin has followed Martin's scheme, except in points where Sarasin definitely states otherwise.

Dudley-Buxton mentioned in his paper ((3), p. 175) that his "measurements were taken in accordance with Martin's definitions," referring no doubt to (14). As I felt doubtful about certain points, Dr G. M. Morant of the Biometric Laboratory, London, at my request very kindly wrote to Mr Buxton and received the following reply in a letter dated Oxford, 5th February, 1927:

"You will find most conveniently my technique summarized in the British Association Anthropometric Investigations in the British Isles, reprinted from the B.A. Report 1908. My measurements are, following their numbering:

A. 1. G.O.L., 2. G.B., 4. M.F.D.
B. 1. U.F.H., 2. T.F.H., 3. Biz.B., 5. E.O.B., 8. Big.B.
C. 1. N.H., 4. N.B.
E. 1. Stature."

Shirokogoroff mentions ((27), p. 1) "the list of measurements elaborated by the

International Commission in Geneva in 1912," and refers to the *Revue anthropologique*, Nos. 7—8, Juillet—Août, 1913. The Geneva Commission (7) does not however give any measurements on the head as its work was restricted to measurements on the body. But as the Geneva Commission was a continuation of the earlier Monaco Commission of 1906 (20) I have assumed that Shirokogoroff really intended to refer to this latter list. According to Shirokogoroff this list " is generally used in Russia," and " the list elaborated by the Commission in England does not differ from this except in the measurements of the physiognomical length of the head." From the fact that names used by Shirokogoroff are usually those of the B.A. Report (21), and that the physiognomical face length does not occur in the latter, I have identified his " list elaborated by the Commission in England " with the B.A. Report (21) and have used it in conjunction with the Monaco standard (20) in connection with Shirokogoroff's measurements.

Systematic definitions of measurements on the head are not given in *Anthropologica Suecica* (13) but notes are given here and there from which it is possible in many cases to reconstruct the definitions fairly satisfactorily.

II. Definitions of Measurements of the Head (Flesh).

(1) *Head Length.*

5. Buxton and Shirokogoroff follow the B.A. Report (21): " No. A 1 Maximum Length. From the most prominent point of the glabella to the most distant point on the back of the head, known as the occipital point." *Notes and Queries* (1874, 1892, 1899, and 1912) and the British Association Schedule (1895) all give the same definition which is in conformity with the corresponding measurement on the skull. Attention is drawn to the necessity of taking the measurements in the median vertical plane. It corresponds to the International Commission No. 1 " Longueur maxima de la tête ou diamètre antéro-postérieur maximum. Même technique que pour le crâne ; ne pas presser" ((20), p. 569). The technique on the skull is described as follows: " C'est le plus grand diamètre dans le plan sagittal et médian du crâne."

" Points anatomiques: en avant: le point le plus saillant de la protubérance intersourcilière (glabelle de Broca); en arrière: le point le plus saillant du susoccipital donné par le maximum d'écartement des branches du compas" ((20), p. 563).

Sarasin presumably follows Martin's definition ((14), p. 157): " No. 1 Grösste Kopflänge (diamètre antéro-postérieur maximum ou glabellaire ; maximum glabello-occipital length): Geradlinige Entfernung der Glabella von Opisthokranion, d. h. von dem am meisten hervorragenden Punkte des Hinterhauptes in der Median-sagittalebene."

Koganei also states definitely ((8), p. 254) that it is the greatest glabello-occipital distance, " die grösste Entfernung zwischen dem Stirnnasenwulst und Hinterhaupt."

Hrdlička defines it as the " maximum glabello-occipital diameter of the vault " ((5), p. 68). The procedure described by him is a little different from the B.A. Report

and Martin, but it does not appear likely to make any appreciable difference in the measurements.

Von Luschan follows Broca's instructions ((1), pp. 105—106): "C'est le plus grand écartement qu'on puisse donner au compas sur la ligne médiane du crâne. L'une des branches est appliquée sans pression au-dessus de la racine du nez, sur le point culminant de la glabella ou bosse nasale; on l'y fixe avec la main gauche pendant que la main droite promène l'autre extrémité du compas sur le derrière de la tête...."

Kubo states that his measurement is the one called "Langsdurchmesser des Schädels" by Schmidt ((25), p. 105): "Der vordere Messpunkt liegt in der Mittellinie der Stirn, dicht über dem Glabellawulst, wo dieser wenig oder gar nicht entwickelt ist, einen kleinen Querfingerbreit über dem Niveau der Augenbrauenbogen." This of course is quite different from the glabello-occipital length, and is really the distance from a point ranging from Glabella to Ophryon and the occipital point. It is, however, very curious that Kubo appears to think that the Glabella and the Ophryon are one and the same point. He says in one place ((10), p. 179), "ich habe...die Entfernung zwischen Ophryon (Glabella) der Stirn und dem äussersten Punkte des Hinterhauptes gemessen." Apart from the use of the Glabella within brackets, evidently as a synonym for Ophryon, he goes on to explain in the very next sentence: "Meine Kopflänge ist nicht die Projektions-, sondern die direkte Kopflänge von der *Glabella* (italics our own) bis zum hervorragendsten Punkte des Hinterhauptes, der durchaus nicht immer die Protuberanz ist" ((10), pp. 179—180). It scarcely needs pointing out how careful one must be in making comparisons.

6. One point requires notice. The B.A. Report explicitly states ((21), p. 8) that "the pressure of the points of the callipers on the head should be as much as can be comfortably borne by the person under examination*." Schmidt too recommends the application of fairly strong pressure: "Die beiden Stangenzirkelarme sind ziemlich stark an das Haupt anzupressen" ((25), p. 105).

On the other hand Broca (1) lays down that the measurement should be taken without pressure ("sans pression"), and von Luschan following Broca definitely states that his measurements were taken without pressure: "ich bei der Bestimmung von Länge und Breite des Kopfes gar nicht drücke" ((12), p. 334). Martin is also definitely of the same opinion: "Werden die Spitzen des Instrumentes unter so starkem Druck, als es das Individuum aushalten kann (Vorschrift der British Association), an die Kopfhaut angepresst, dann wird das Mass zu klein. Eine derartige Messung ist aber nicht nur schmerzhaft, sondern auch ungenau und daher zu verwerfen" ((14), p. 157). He adds in a footnote on the same page: "Ein Vergleich der Kopfmasse mit den Schädelmassen ist natürlich nur dann möglich, wenn nach absolut gleicher Technik gemessen und das Instrument nicht in die Kopfhaut eingepresst wird."

* *Notes and Queries* (4th edition, 1912) practically follows the B.A. Report and says "the pressure between the points of the callipers should be firm, but not uncomfortable" (p. 6). Earlier British authorities are silent on this point.

The Monaco Commission distinctly says: "ne pas presser" ((20), p. 57). Hrdlička translates it as "do not press" ((5), p. 21), but curiously enough adds the following note, "a moderate amount of pressure is of course necessary; the instruction is directed against hard pressure." I do not understand the force of "of course," nor what authority Hrdlička has for interpreting the International Instructions in the way he does. It is however clear that Hrdlička's measurements were taken with a moderate degree of pressure.

Shirokogoroff does not appear to be aware of the direct opposition between the two authorities cited by him: the B.A. Report (21) and the Monaco Agreement (20), and it is uncertain what procedure he followed. The fact that he relegates the B.A. Report to a footnote, while he refers to the International Commission in his text, may be interpreted as implying that he attached greater importance to the latter. If this interpretation be correct then his measurements would be "without pressure." This is supported by the fact that he gives the Korean Head Length and the Head Breadth as 183·9 mm. and 153·7 mm. respectively against 181·4 mm. and 150·9 mm. given by Kubo. This amounts to an excess of 2·5 mm. and 2·8 mm. respectively for measurements taken without pressure*.

7. The Maltese (Buxton), the Korean (Kubo), the Egyptian and the Old American (Hrdlička) and very likely the Aino (Koganei) measurements were taken with the application of a certain amount of pressure and are not strictly speaking comparable with the measurements on the Cretans (von Luschan), the New Caledonians (Sarasin) and the Mongolian people of the far east (Shirokogoroff) which were taken without pressure. Definite information is not available about the Swedish data, but assuming that no pressure was employed they would belong to the second group.

Thus we find that Maltese, Aino, Egyptians and Old Americans are comparable with one another, as also Cretans, New Caledonians and the Mongolian and possibly the Swedish samples, but members of one group are not comparable with those belonging to the other group. The Korean Head Length stands apart and is not comparable with any other.

(2) *Head Breadth.*

8. B.A. Report ((21), p. 8): "Maximum Breadth. Measured wherever it can be found above the plane of the ear-holes. The callipers should be held in a vertical transverse plane and moved about until the maximum diameter is ascertained, the observer being careful to keep the points of the callipers exactly opposite to one another." B.A. Schedule ((24), p. 3) states that the maximum breadth is to be found "usually about the top of the ears," but does not otherwise restrict the location. *Notes and Queries* ((17), 2nd edition, 1892) states that it is to be taken "wherever it may be except low down behind the ears."

Martin ((14), No. 3, p. 159): "Grösste Kopfbreite (diamètre transversal maximum). Geradlinige Entfernung der beiden Eurya voneinander, d. h. grösste Breite

* With a mean standard deviation of about 7·0 mm. a difference of 3 mm. in one of the Head Lengths (due to difference of technique) will make a difference in $(M - M')^2$ of nearly 9 times the variance for samples of 100, and will be definitely appreciable.

senkrecht zur Median-sagittal-Ebene, wo sie sich findet....Die Messpunkte müssen in einer Horizontal- und Frontal-Ebene liegen."

Broca ((1), p. 166) mentions "Les extrémités de ce diamètre sont en général situées à deux ou trois travers de doigt en arrière et au-dessus du bord supérieur de l'oreille ; mais cela est loin d'être constant, et il faut toujours tâtonner un peu pour trouver le véritable maximum du diamètre transversal." Topinard ((30), p. 36) too mentions "above the ears," but does not otherwise restrict the location.

The International Commission No. 3: "Largeur maxima de la tête ou diamètre transverse maximum" ((20), p. 569). The technique is the same as on the crania: "C'est le plus grand diamètre horizontal et transversal qu'on puisse trouver avec le compas d'épaisseur sur la boîte cranienne. Point anatomique, déterminé seulement par le maximum, mais si ce dernier tombait sur les crêtes sous-temporales, il faudrait éviter leur saillie, en plaçant le compas au-dessus" ((20), p. 563).

Schmidt ((25), No. 31, "Breite des Schädels," p. 105) lays emphasis on the need for both arms of the instrument lying in the same horizontal plane, but does not otherwise discuss the question of the position where the measurement is to be taken.

Hrdlička ((5), p. 96) describes it as the "maximum breadth of the skull above the supramastoid and zygomatic crests."

As regards the location of the end-points there is substantial agreement, although all the definitions are not rigidly identical. In particular cases, especially for deformed individuals, differences may arise owing to the slight ambiguity in the description of the region where this measurement is to be taken.

9. As regards pressure, B.A. Report ((21), p. 8) makes the same remarks as in the case of Head Length: "the pressure on the points of the head should be as much as can be comfortably borne by the person under examination." Other authorities do not mention anything explicitly on this point; but it may be presumed that they would adopt respectively the same procedure as in the case of Head Length. On this assumption, Maltese, Korean, Egyptian, Old American, and very likely Aino data would form one group, and Cretan, New Caledonian, and Mongolian a second one, comparison being possible within either group. It is probable that the Swedish measurements also belong to the second group*.

(3) *Cephalic Index.*

10. Maltese, Aino, Egyptians, and Old Americans may be compared with one another, so also Cretans, New Caledonians and Mongolians and possibly the Swedish. The Korean index is not comparable with any other, even apart from considerations of "pressure"†.

* With a mean S.D. of 5·72 mm. approximately, a difference of 3 mm. in Head-Breadth (which may easily arise owing to the application of pressure) will make a difference of about 14 times the variance in $(M - M')^2$ in samples of 100.

† If the same observer uses the same amount of pressure in taking both the length and breadth measurements of the head, then the error in the cephalic index will be partly compensated, and will not generally be appreciable in samples of small or moderate size, but may become significant for large samples.

(4) *Height of the Head.*

11. Koganei defines this measurement ((8), p. 254) as "die senkrechte Höhe vom oberen Rande der äusseren Gehöröffnung bis zum Scheitel." Kubo gives an identical definition ((10), p. 223), and his measurements are therefore comparable with those of Koganei's. This is also* the definition accepted by the Monaco Commission, No. 3: "Hauteur de la tête (placée bien d'aplomb sur ses condyles)....Points anatomiques: en haut: vertex; en bas: bord supérieur du trou auditif, dont le point de repère (toujours à vérifier) est ordinairement le fond de l'échancrure comprise entre le tragus et l'hélix" ((20), p. 570).

Shirokogoroff adopts a different procedure. He explains that he measured the total stature as well as the height of the ear-hole from the ground. "It was thus possible to work out the height of the head by subtraction of the height of the ear-hole from the stature" ((27), p. 2). It is not clear which portion of the ear-hole (the upper border, or the centre, or the lower border) he took as his end-point. The usual definition of "height to ear-hole" extends to the *middle* of the ear-hole (e.g. *Notes and Queries*, 4th edition, 1912, p. 4). Assuming that Shirokogoroff also measured to the middle of the ear-hole, his measurements of the Height of the Head would not be comparable with those of Kubo and Koganei.

Hrdlička in his monograph on the Egyptians ((4), p. 48) mentioned that he took the measurement "by a spreading and sliding compass. The branches...are introduced well into the auditory meati and allowed to rest on the floor. The expansion of the instrument is noted, with the scale held over the bregma region; the distance from the bregma region to the lower edge of the scale is measured by the rod of the 'compas glissière,' and a simple arithmetical process gives the biauricular line-bregma height." He explains that this gives the distance between "the line connecting the floor of the auditory canals to the scalp over the head."

In his *Anthropometry* ((5), p. 21), in a footnote to his translation of the definition given by the Monaco Commission, Hrdlička says "the height from the middle of the line connecting the floor of the external auditory canals to bregma is now more in vogue," and on p. 70 states that this method "has been practised by the author since 1898." His measurements are therefore not comparable with those of either Koganei and Kubo, or Shirokogoroff.

Hrdlička mentions that his method "gives results somewhat higher than those obtained by Gray's radiometer" ((4), p. 48). In fact for the same individual, Hrdlička's method would give the greatest, Kubo and Koganei's the lowest readings, while Shirokogoroff's value would lie between the two.

To sum up, Kubo's and Koganei's measurements are comparable with each other but with no other. Nor are the other measurements comparable among themselves†.

* [I am not clear that this is so, the "vertex" is the highest point of the head above the Frankfurt plane, it is not necessarily identical with the "apex," the point in the median plane vertically above the biauricular line. From biauricular line to vertex will usually be greater than the vertical height of the head. K.P.]

† The difference between Kubo and Shirokogoroff for the Korean sample is 7·2 mm. With a mean S.D. of 6·13 mm. this will make a difference of more than 125 times the variance in $(M - M')^2$. Hrdlička mentions in one place ((5), p. 72) that his method "with due care...gives results which may vary within less than 3 mm." Assuming that a difference of even 3 mm. is made by adopting different procedures, a difference of more than 12 times the variance will be made in $(M - M')^2$ for samples of 100.

(5) *Horizontal Circumference of the Head.*

12. Koganei describes it ((8), p. 255) as the circumference measured over the Glabella and the furthest part on the back of the head ("über der Glabella und dem vorspringendsten Punkte des Hinterhauptes gemessen"). This corresponds to No. 45, Horizontalumfang des Kopfes, of Martin ((14), p. 173) which passes over the Glabella and the "vorspringendsten Punkte des Hinterkopfes (Opisthokranion)," and to the B.A. Report No. A 4 Maximum Circumference ((21), p. 8) which is described as the circumference "measured by passing the tape horizontally (*sic*) round the cranium at the level of the glabella in front and the occipital point behind."

Kubo adopts the same definition as Koganei, and says ((10), p. 212): "Ich habe den Umfang desselben über der Glabella und dem vorstehenden Punkte des Hinterhauptes so gemessen, dass das Bandmass genau in einer transversalen Ebene, nicht auf der einen Seite höher als auf der anderen zu liegen kam." But we have already seen that he believes the Glabella and the Ophryon to be identical points. This confusion makes his present definition a little uncertain. If he has actually measured it over the Ophryon then his circumference would correspond to Martin's ((14), p. 173) "No. 45 *a* Horizontalumfang rund um den Kopf, jedoch über das Ophryon (statt über die Glabella)."

13. The qualifying phrase "horizontal" appears to be a little misleading. Broca says ((1), p. 277): "Elle n'est pas horizontale, comme son nom pourrait le faire croire, mais toujours plus ou moins oblique, attendu qu'elle passe plus haut en avant qu'en arrière....On n'oubliera pas que cette courbe est un maximum."

(6) *Nasal Height.*

14. The Nasal Height (which is a constituent of the Nasal Index) is recognised to be one of the most important measurements in cephalometry, and yet the most diverse methods appear to have been used in measuring it. Martin ((14), pp. 166—167) gives us no less than five different definitions under the general name "nasal height" or "nasal length." The upper point of reference has been taken as (i) the Ophryon, (ii) the intersection (in the median-plane) of the tangent line to the two orbital edges, (iii) the nasion (or what is believed to be the point on the flesh corresponding to the naso-frontal suture on the skull), or (iv) the deepest part of the nasal bridge. There is greater agreement about the lower point of reference; it has usually been taken as the sub-nasal point (where the nasal septum joins the upper lip). But in at least one case (Shirokogoroff) this latter point seems to have been confused with the pro-nasal point or tip of the nose. The surprising thing in this connection is the lack of appreciation on the part of eminent anthropometrists of the need for noting a precise description of the definition actually adopted in practice.

B.A. Report ((21), p. 9) gives: "No. C 1 Nasal Height. From the nasion to the sub-nasal point."

Martin ((14), p. 166): "No. 21 Höhe der Nase (fälschlich Länge der Nase, Nasobasallänge, hauteur ou longueur du nez). Geradlinige Entfernung des Nasion vom Sub-nasale."

Broca ((1), p. 182): "No. 4 Longueur du nez. Elle se mesure en appliquant transversalement les deux branches de la glissière sur la racine du nez, et sur le point sous-nasal."

Topinard ((30), p. 357): "The height is taken with the sliding compass, vertically from the root to the base of the nose, as on the skull."

The Monaco Commission ((20), p. 570) gives No. 12 "Hauteur du nez. Points anatomiques: en haut: nasion; en bas: sous-cloison du nez, au niveau de son union avec la lèvre supérieure. Ne pas presser."

Hrdlička himself ((5), p. 74) has a practically identical definition: "Nose: Length.—The length (or 'height') of the nose from the nasal septum where this joins the upper lip, to the nasion."

Koganei ((8), p. 255) calls it "Nasenlänge" and defines it as "die Entfernung der Nasenwurzel vom Ansatze der Nasenschiedewand an der Oberlippe." Kubo calls it ((10), p. 317, No. 34) "Höhe (Länge) der Nase," gives an identical definition and adds "die Nasenlänge ist also nicht auf dem Rücken, sondern an der Basis der Nase, zwischen Nasenwurzel und Nasen-Oberlippen-Winkel, zu messen," which is practically a quotation from Schmidt ((25), p. 108, No. 49). This makes it clear that the sub-nasal point and not the tip of the nose is the lower point of reference.

15. It will be noticed that in the above definitions the upper point of reference is given as the "nasion," the "Nasenwurzel" or the "root of the nose," and in spite of the verbal agreement this is where trouble begins, because there exists a complete lack of agreement about the precise definition of these terms.

B.A. Report ((21), p. 6) defines the "Nasion" as "the bottom or deepest part of the depression between the forehead (Glabella) and the nose, or, in other words the most depressed part at the root of the nose*." Other British authorities are in agreement on this point. *Notes and Queries* ((19), 1912, p. 7, No. 7) gives "Nasal Height—from the nasion or the most depressed point at the root of the nose to the sub-nasal point or angle between the septum of the nose (i.e. the partition between the nostrils) and the upper lip." *Notes and Queries* ((18), 1899, p. 22) gives the following description: "...the upper point at the termination or root of the nose between the eyes....This last point is sometimes a little difficult to determine. There is a small transverse fold of the skin (sometimes two folds) at the root of the nose; it is on this fold, or when there are two folds, between the folds, that the upper point of the instrument should rest, generally about two millimetres above the level of the transverse axis of the eye." This description is a verbatim copy from the second edition (1892) of the same book. The Brit. Assoc. Schedule ((24), 1895) describes the nasal height as the distance "from the furrow at root of nose to the angle between the nose and the upper lip in the middle line."

The British view therefore identifies, and has always identified, the "Nasion" with the root of the nose, defining the latter term as the most depressed part of the nose where the skin has frequently a fold or two.

* It is curious however that the illustration on p. 387 (original report) shows the nasion on the flesh as coinciding with the naso-frontal suture on the skull. This is one example of the lack of preciseness in essential points found in existing standards.

16. Other authorities are however of quite different opinions. Martin is quite explicit on this point ((14), p. 129): "Die Nasenwurzel entspricht nicht der am tiefsten eingesattelten Stelle des Nasenrückens, die meist im oberen Teil der Ossa nasalia gelegen ist, sondern der Sutura naso-frontalis, deren Verlauf nach einiger Uebung trotz des vorhandenen Nahtgewebes und des meist dünnen M. procerus (M. depressor glabellae n. H. Virchow) auch am Lebenden festgestellt werden kann. Man findet den Punkt am besten, wenn man seine rechte Hand ruhig auf den Kopf des zu messenden Individuum legt und mit dem lateralen Rand der Daumenbeere unter leichtem Druck die Haut auf der Nasenwurzel auf und ab schiebt. Man beachte, dass das Nasion in der Regel im Niveau der medialen Enden der härenen Augenbrauen, meist an deren Unterrand, nicht in der Höhe der Lidspalte gelegen ist." He concludes by observing: "Von manchen Autoren wird fälschlich die am tiefsten eingesattelte Stelle der Nase als 'Nasenwurzel' bezeichnet."

The illustration in his book ((14), fig. 38, p. 128) clearly shows that according to Martin the nasion is situated at a considerable distance above the most depressed part of the nasal bridge. In fact Martin had stated at a different place ((15), pp. 393—394): "Allerdings liegt die tiefste Einsattelung der Nase gewöhnlich nicht an der Nasenwurzel, sondern beträchtlich unterhalb derselben; ich habe z. B. bei Senoi von Ulu Gopei eine Entfernung von 7 bis 9 mm. zwischen diesen beiden Punkten feststellen können."

The general tendency among continental authorities seems to be to make the upper point of reference on the flesh correspond to the nasion on the skull, which is defined as the naso-frontal suture. For example, Broca says ((1), p. 139): "La dépression transversale qui la sépare du nez s'appelle la *racine du nez*, ou *point nasal* ou *nasion*; elle correspond sur le squelette à la suture qui unit l'os frontal aux os nasaux." Topinard also had definitely used the words ((30), p. 357) "as on the skull," the implication clearly being some kind of correspondence between the points of reference on the flesh and on the skull.

Hrdlička ((5), p. 72) says: "The nasion should correspond as closely as possible to the anatomical nasion, i.e. the mid point of the naso-frontal suture."

Von Luschan ((11), p. 41) states: "Man fühlt stets eine kleine Vertiefung die genau der Naht zwischen Nasenbein und Stirnbein entspricht." He also notes: "Laien fehlen regelmässig dadurch, dass sie die Nasenwurzel wesentlich tiefer suchen und etwa die Mitte der Linie zwischen den inneren Augenwinkeln zum Ausgangspunkt ihrer Messung machen wollen."

17. We thus find two sharply contrasted views about the location of the nasion, and hence about the upper point of reference in taking measurements of the nasal height. There is general agreement about the fact that the nasion on the skull is usually situated considerably higher up than the most depressed part at the root of nose on the the flesh. We have already seen that Martin found the difference as high as from 7 to 9 mm. in certain individuals. The anatomical work of R. Havelock Charles ((2), 1894) on 54 male and 8 female cadavers shows that the nasal height on the skull is almost certainly longer than the nasal height on the living. But

unfortunately Havelock Charles does not provide us with any statement of how he measured the nasal height on the cadaver. He talks of the visible "root" of the nose, but is the "root" of the nose the most receding point of the nasal bridge, when the head is adjusted to the Frankfurt horizontal, let us say? He does not write as if he had recognised that the height of the nose on the living had been measured in a variety of ways. When we come to Havelock Charles's *lower* point of measurement we do not find that he has stated in the case of either skull or cadaver what he has taken for it. All we are told is that "Having carefully measured with sliding compasses, the nasal diameters of 62 'subjects,' I removed the integuments, etc., and having cleared the naso-frontal suture and anterior nasal aperture, I again took the diameter*." He does not tell us whether, in measurement on the subject, he took the *sub-nasale*, or the lowest point of the fleshy aperture, or whether if the former he applied pressure or not. Nor again are we told whether in measuring the nasal height on the skull he took as his lower point, the lowest point of the pyriform aperture, the base of the nasal spine, or the spine itself†. However, Charles found the nose in the flesh to be less than the nose on the skull in height by -2 to $+16$ mm. It does not follow that this difference is due to the "root of the nose" not being the nasion. A portion of it depends on the depth of the *sub-nasale* being in the flesh below the nasal spine. Thus it seems probable that the difference between the two lengths is only a minimum difference of the distance from the nasion to the "root of the nose," whatever the latter exactly connotes. We have for the 54 Panjabi males, Mean Difference = 4·33 mm. $\pm$ ·32, and for the eight (!) Panjabi females, 3·50 mm. $\pm$ ·26. Thus it seems probable that the distance from nasion to root of nose is at least 4·5 mm. for males and 3·5 mm. for females of this race, and it may vary considerably with other races. As Havelock Charles emphasises nasal indices thus measured on the living and on the skull are not comparable‡.

18. It is a pity that Charles did not make any attempt to find out how accurately a point on the flesh could be located to correspond with the naso-frontal suture. To a layman it certainly appears doubtful whether this point can actually be determined with any degree of reliability. Hrdlička ((5), p. 72) says in this connection: "In a certain portion of subjects this point may be felt by the observer's finger nail or the point of a pencil§." But Hrdlička immediately remarks: "but in a majority we must rely on knowledge of its location derived from extensive observation on skulls and dissecting room material. It is always situated above a horizontal line connecting the two inner canthi." Thus in a *majority of cases* (according to Hrdlička's own admission) the location of the point must be made by what is nothing else but guess-work, and in case of observers without any experience of the dissecting room, by perfectly random guess-work without any basis of previous knowledge. From the

* *loc. cit.* p. 2. Havelock Charles without definition uses the symbol μ, customary for microns, for millimetres.

† He may have taken it again to the sub-nasal point, but this is not usual in measuring the nasal height of the skull.

‡ A difference of 4·0 mm. in the mean will make a difference of more than 56 times the variance in $(M - M')^2$ for samples of 100 (with a mean S.D. of about 3·7 mm.).

§ This is also the course recommended by Martin.

point of view of metrology an external or surface landmark, such as the most depressed part of the root of the nose, would appear to be far more satisfactory.

19. Buxton (Maltese) followed the B.A. Report, and must therefore have adopted "the most depressed part at the root of the nose" as his upper point of reference*.

Sarasin (New Caledonians) also has taken the same definitions. In his own work on *Versuch einer Anthropologie der Insel Celebes* (Wiesbaden, 1906), *Zweiter Teil*, p. 29, he says, referring to this question of the location of the nasion: "Ich befinde mich hier in einem Widerspruch mit Martin, welcher am Lebenden als Ausgangspunkt die Stirn-Nasenbein-Sutur auszutasten sucht; diese liegt aber gewöhnlich beträchtlich oberhalb der tiefsten Einsattelungsstelle der Nase. Ich halte das nicht für richtig, denn das physiognomische Nasenbild des lebenden Menschen wird, wie auch Martin ((15), p. 394) zugibt, die Einsattelungsstelle bedingt und nicht durch die Sutur am Schädel, und dieses Bild der lebenden Nase ist es, was wir durch einen Index ausdrücken wollen. Auch hat es gar keinen Sinn, ein aus der Schädelanatomie gewonnenes Maass mit einem reinen Fleischteilmaass, wie es die Breite der Nasenflügel ist, miteinander in einem Index zu vergleichen." In his work on the New Caledonians ((22), p. 95) he says he has: "als obere Messpunkt für die Nasenhöhe die tiefste Einsattelungsstelle der Nase gewählt, um einen besseren Ausdruck für das Nasenbild des Lebenden zu gewissen."

It is however not clear what procedure Kubo (Koreans) has followed. Kubo constantly refers to Schmidt's book (25) but unfortunately Schmidt is not very explicit on this point. Schmidt ((25), p. 107, No. 43) describes the "Nasenwurzel" as "die tiefste Stelle der Einsattelung zwischen Stirn und Nase," which apparently identifies the root of the nose with the most depressed part of the nose, but continues in the very next sentence to say: "Man fühlt hier bei feiner Haut die Knochensutur durch, welche die Nasenbeine vom Stirnbein trennt." Which renders things quite inconclusive. Kubo also referred with approval to the following note by Baelz ((10), p. 253) about the difficulty of locating the root of the nose in east-Asiatic peoples: "Die Nasenwurzel hat beim Japaner meist eine andere relative Lage, als beim Europäer wir verstehen die tiefste Stelle des Nasensattels. Beim Europäer ist diese Stelle nur wenig unterhalb der Verbindungsstelle beider Augenbrauen gelegen, beim Japaner dagegen ist der Unterschied sehr bedeutend, beim Europäer bildet der Nasensattel eine scharf markierte Stelle, eine Art Winkel, beim Japaner dagegen einen flachen Bogen†." This passage would suggest that Kubo was really seeking for the most depressed part at the root of the nose, and was therefore pointing out the difficulties encountered in locating this point for east-Asiatic peoples. On this interpretation his measurements would be comparable with those of Buxton and Sarasin but the legitimacy of such a comparison cannot be established with certainty.

* Buxton appears to be unaware of or has neglected the divergence between the two authorities cited by him on different occasions, namely in (3) he cites Martin (14), but in the letter to Dr Morant already quoted on p. 3 above he cites the B.A. Report (21).

† For a careful study of the form of the nose see E. von Eickstedt (32).

F. von Luschan (Cretans) and Hrdlička (Egyptians, Old Americans) on the other hand definitely adopted a different definition of the nasion on the living and attempted to locate a point on the flesh which corresponded as closely as possible to the naso-frontal suture on the skull.

It is possible that Koganei (Aino) also did the same thing, for although he does not mention anything explicitly in his paper on the Aino, in a later monograph on the Northern Chinese Soldiers ((9), p. 137) he distinctly states: "Als oberer Messpunkt wurde nicht die tiefste Stelle der Einsattelung zwischen Stirn und Nase, welche wenigstens bei ostasiatischen Völkern gewöhnlich gar nicht der Stirn-Nasennaht entspricht, sondern einige mm. oberhalb derselben genommen, wo eben die Stirn-Nasennaht gelegen ist welche aber nur schwer durchzufühlen ist."

For the Swedish data Lundborg and Linders mention Martin (14) in one place, and it is probable that Martin's definition of the nasion was followed in this case. This hypothesis is supported by the following remarks ((13), p. 11): "...The examiners often cannot locate the nasion (sutura naso-frontalis) with certainty." In footnote 8, on p. 159, commenting on the Morphological Face Height we have: "Hilden measured from the deepest part of the nasal root to the gnathion." Comparing Hilden's mean Morphological Face Height 123·1 (71 individuals) with Lundborg's 126·6 mm. (for 47,387 men) I find a difference of 3·5 mm. in favour of the latter, which supports the conjecture that the latter were trying to identify the nasion on the flesh with the naso-frontal suture on the skull*. There is another footnote, No. 10 on p. 159, "Bryn located the nasion as a point 2 mm. above the deepest part of the nose†." These footnotes suggest that Lundborg and Linders' point of reference was different from those of Hilden and Bryn.

20. We must now consider Shirokogoroff's measurements. He at first uses the term "Height of the Nose" ((27), p. 1), but subsequently adopts "Length of the nose" and continues to use the latter name in most places. He is apparently not aware of the fact that these two names refer to two different entities. The "Nasal Length" is defined by the B.A. Report ((21), No. C. 3, p. 10) and Martin ((14), No. 23, p. 167) as the distance from the nasion to the tip of the nose, and is quite different from the "nasal height" which is measured to the sub-nasal point.

A comparison of Shirokogoroff's means with those of other authors shows that it is improbable that he actually measured the nasal height. For a series of 141 Koreans, Shirokogoroff gives a mean "nasal height" of 40·8 ± ·18 mm., while Kubo's value is 49·0 ± ·09 mm. for 477 individuals. This gives a difference of 8·2 ± ·31, or a difference of about 26·5 times its own probable error.

Again for Northern Chinese we have Koganei's value ((9), p. 144) of 53·0 mm. for 968 individuals against Shirokogoroff's of 41·8 for 391 individuals, giving a difference of 11·2 ± 0·15 mm. or a difference of 80 times the probable error. Such

* I should note however that Hilden's measurements referred to the inhabitants of the island of Rüno, which is *believed* to be populated by Swedes, while Lundborg and Linders' measurements were all taken on persons of authentic Swedish extraction. So that we cannot be absolutely certain that the difference of 3·5 mm. was due to a difference in technique and not to a difference of race.

† This footnote is interesting as giving still another definition of the nasion on the living.

huge differences are, of course, extremely unlikely to occur in different samples of the same population, and they strongly suggest some fundamental differences in the technique. In fact for the two Korean series, for measurements which are known to be comparable, the next biggest difference is 0·81 mm. only for the Bigonial Diameter which is just over 5·2 times the standard deviation; a big difference no doubt, but still of quite another order from the difference in nasal height.

It is unlikely that differences as large as 8·2 and 11·2 mm. are due solely to a difference in the location of the upper point of reference, though that must be admitted as a possible explanation. The mean Panjabi difference found by Havelock Charles is 4·2 mm., but the form of the nasal bridge varies widely from race to race and the corresponding Chinese value may be much greater.

The absolute magnitudes of Shirokogoroff's "nasal height" are suspiciously low. For example for Koreans it is 40·8 mm. For certain other Chinese samples it is still lower, e.g. Kwantung (total) 40·0 mm., Kwantung (non-selected) 39·3 mm. The lowest value of the nasal height which I could find in standard literature was 41·7 mm. for Kachin in Martin ((14), p. 451).

21. The natural explanation suggested by Shirokogoroff's frequent use of the term "length of the nose" would be that he had actually taken his measurements to the tip of the nose (instead of to the sub-nasal point) as his lower point of reference. This explanation is rendered all the more plausible by the peculiar history of the two names.

In both the 2nd and 3rd editions of *Notes and Queries* published in 1892 and 1899 respectively, as also in the B.A. Schedule of 1895, the names "Length of the Nose" and "Nasal Length" (*and not "Nasal Height"*) are used to denote the distance from the nasion to the sub-nasal point. Among British authorities, the name "nasal height" was used in its modern sense for the first time in the British Association Report of 1908. At the same time the name "nasal length" was also defined for the first time to denote the distance to the tip of the nose. Thus among British authorities "nasal length" would have one meaning up to 1908, and an altogether different meaning since that year. The same anomalous position had existed also in continental literature. In fact in earlier years the general tendency everywhere was to call the distance between the nasion and the sub-nasal point (and *not* the tip of the nose) the "nasal length." For example Broca ((1), p. 182) in 1879 calls it "Longueur du nez," and Schmidt ((25), p. 108) in 1888 calls it "Länge der Nase." On the continent, the International Commission of Monaco of 1906 seems to have been the first authoritative body to introduce the modern name "nasal height" for this particular measurement.

Koganei writing in 1893 naturally uses "Nasenlänge," while Kubo in 1913 calls it "Höhe der Nase," but adds the word "Länge," within brackets after "Höhe," evidently to link it up with the older name. Von Luschan in 1913, Sarasin in 1916, Buxton in 1922, and Lundborg and Linders in 1926 all use the modern name "nasal height." Curiously enough Hrdlička continued to use the older name "length" even in 1920, but he added the word "height" within brackets ((5), p. 74).

22. Summing up we find that the data considered here fall into three distinct groups:

(i) Maltese, New Caledonians, and possibly Koreans, measured according to the British definition of the most depressed part of the nose as the upper point of reference.

(ii) Cretans, Egyptians and Old Americans, and possibly the Aino and the Swedish data, measured according to a definition which makes the upper point on the flesh correspond as closely as possible to the naso-frontal suture on the skull.

(iii) The Mongolian data measured very likely to the tip of the nose instead of to the sub-nasal point.

Members of each group are probably comparable with one another, but not with members of a different group.

(7) *Nasal Breadth.*

23. B.A. Report ((21), p. 10): No. C 2 "Nasal breadth. The greatest diameter, measured without pressure, between the wings of the nose." This may be identified with Martin's ((14), p. 162) No. 13 "Breite der Nase (Untere Nasenbreite, Nasenflügelbreite): Geradlinige Entfernung der beiden Alaria, d. h. der Punkte der grössten seitlichen Ausladung der beiden Nasenflügel voneinander." Martin also mentions that the measurements must be taken without pressure ("ohne den geringsten Druck auf die Nasenflügel"). Topinard had given the definition "the breadth from the widest portions of the alae" ((30), p. 357), and the Monaco Commission of 1906: "No. 13 Largeur du nez. Points anatomiques: face externe des ailes du nez" ((20), p. 571), and Hrdlička ((5), p. 74) agree with the above definition.

Measurements taken by Buxton (Maltese), Hrdlička (Egyptians, Old Americans). Shirokogoroff (Mongolians), Sarasin (New Caledonians), and von Luschan (Cretans) are therefore probably comparable with one another.

Koganei (Aino) and Kubo (Koreans) on the other hand call it "Nasenbreite" ((8), p. 255, No. 15) and "Breite der Nase" ((10), p. 323, No. 35) respectively, and both define it as "der Abstand der Ansätze der Nasenflügel voneinander." The use of the word "Ansatz" (onset, beginning) shows that this measurement is the same as Schmidt's "Breite der Nase" ((25), p. 108, No. 50): "an der Ansatzstelle der Nasenflügel zu messen," which may also be identified with Martin's ((14), p. 163) No. 13*a*: "Grösste Breite am Hinterrand der Nasenflügel, da, wo sie an der Wangenhaut festgewachsen sind." Koganei and Kubo's measurements while comparable with each other are therefore not comparable with the first group*.

(8) *Nasal Index.*

24. Comparing the classification under (6) Nasal Height and (7) Nasal Breadth we get the following distinct groups for the Nasal Index:

(i) Maltese, New Caledonians.

(ii) Koreans.

* A difference of only 5 mm. will make a difference in ($M - M'$) of 12 times the S.D. in samples of 100 (with mean S.D. = 2·9 mm.).

(iii) Cretans, Egyptians, Old Americans.

(iv) Aino.

(v) Mongolians.

Thus for such an important characteristic as the "Nasal Index" valid comparison is not possible for the greater part of the data*.

It has been shown above (Sections 6—8) that the nasal measurements of the living have been determined in a variety of different ways by different schools of anthropologists. The divergences dependent on that fact are evidently large compared with the inter-racial variations of the measurements, and there is no satisfactory method of adjusting the data so that they may be compared directly. The value of these statistics is greatly lessened owing to the deplorable lack of agreement between leading anthropologists of different countries. The nasal index of both the skull and the living head has been recognised as a character which should be of great importance in distinguishing different races of man, and various attempts have been made to assemble the comparative material available. The earliest list of any permanent value appears to be that furnished by Topinard in 1885 ((29), p. 303). No references are given, but all the racial means may have been determined in accordance with the French technique originated by Broca. Two years later Collignon published a more extended list, again without references†. It included means provided by Rudolf Virchow. The best known tables of living nasal indices of more recent date are those of Martin ((14), pp. 447—449), Thomson and Dudley-Buxton‡ and Deniker§. All agree in giving no references whatever. Of the three Deniker is the only one who comments on the fact that different methods of technique were used in determining the measurements. Martin's list includes data provided by English (Duckworth, Beddoe, Thurston, Risley‖, Haddon and the Census of India‖), French (Collignon, Deniker, Legendre, Mondière and Chantre), and German (Schmidt, von Luschan, Fischer, K. Ranke and Martin) anthropologists and many others. Of the series compared in detail above Koganei's Aino, Hrdlička's Egyptian and von Luschan's Cretan means are quoted, and it was shown there that no valid comparison can be made between the Aino measurements on the one hand and the Egyptian and Cretan on the other. It would be easy to find other groups which cannot be legitimately compared. Thomson and Dudley-Buxton cite neither

* Differences due to technique will be appreciable even in small samples. For example, with a mean S.D. of 7·97, a difference of 5 units in means will made a difference in $(M - M')$ of 4·4 times the S.D. of the difference in samples of 100.

† R. Collignon: La Nomenclature quinaire de l'Indice nasale du vivant. *Revue d'Anthropologie*, 6ième année, 3ième série, T. 2, 1887, pp. 8—19.

‡ Arthur Thomson and L. H. Dudley-Buxton: Man's Nasal Index in Relation to Certain Climatic Conditions. *Journal of the Royal Anthropological Institute*, Vol. LIII, 1923, pp. 92—122. The living nasal indices are given in Appendix I (pp. 116—119).

§ J. Deniker: *Les Races et les Peuples de la Terre*, 2nd edition, Paris, 1926. The living nasal indices are given in Appendix III (pp. 718—721).

‖ The Census of India data were obtained under Risley's directions and it is probable that the French technique (as defined by Topinard) was used throughout.

references nor authorities for the material which they quote, but they certainly use the means given in the Census of India* and by Deniker, Delisle and Collignon. Possibly all the nasal indices given were found in accordance with the usual French technique, but the omission to supply any information relating to this vital matter is a regrettable one. Deniker (*op. cit.* pp. 98—99) comments on the fact that German and French anthropologists have determined nasal measurements in different ways and he says that all the means given in his list "ont été prises d'après la méthode Broca-Collignon." On turning to his table, however, we find several values given by E. Schmidt, Fischer and Sarasin. Of the series compared in detail above, Sarasin's New Caledonian, Hrdlička's Egyptian and von Luschan's Cretan means are given. It has been shown that the comparison of the first with the last two is quite invalid. We are obliged to conclude that no single one of these lists can be considered to have any statistical value until it has been carefully revised or confirmed. It is extraordinary that the anthropologists who prepared them should have been so oblivious to the demands of empirical science.

(9) *Nasal Depth.*

25. B.A. Report ((21), p. 10): "No. C 2 Nasal depth. From the sub-nasal point to the most projecting point of the tip of the nose." Martin ((14), p. 167) gives practically an identical definition—No. 22 "Länge des Nasenbodens (Tiefe der Nase, fälschlich Höhe der Nase, Nasenelevation; saillie de la base du nez; nasal depth): Projektivische Entfernung des Subnasale vom Pronasale." The Monaco Commission, No. 14 "Saillie de la base du nez," also gives the same definition ((20), p. 571): "Points anatomiques en avant: le point le plus saillant du lobule nasal; en arrière: le point où le plan médian est coupé par la ligne transversale joignant le point le plus reculé de chacun des plis naso-labiaux. Prendre la distance en projection de ses deux points avec un instrument approprié."

Both Koganei ((8), p. 255, No. 13) and Kubo ((10), p. 331, No. 36) define it as "die Prominenz der Nase vom Gesicht, gemessen vom Ansatze der Nasenschiedewand an der Oberlippe bis zur Spitze," and their measurements are comparable.

26. We again notice a terminological ambiguity. Koganei calls this measurement "Nasenhöhe," while Kubo calls it "Tiefe der Nase," but adds the word "Höhe" within brackets.

(10) *Morphological (Total) Face Height.*

27. B.A. Report ((21), p. 9): "No. B 2 Total Face Height. From the nasion to the lower edge of the point of the chin. A contact measurement†." Buxton calls it the "total facial height‡." The Monaco Commission, No. 9 "Diamètre naso-

* The Census of India data were obtained under Risley's directions and it is probable that the French technique (as defined by Topinard) was used throughout.

† A contact measurement is one to be taken without pressure.

‡ I may note in passing that this same measurement is called "Length of Face" in *Notes and Queries*, 1874 ((16), p. 4, No. 41), 1892 ((17), p. 24) and 1899 ((18), p. 24). But the name is changed to "total facial height" in the 1912 edition ((19), p. 7).

mentonnier*. Même technique que sur le squelette, en pressant un peu comme pour le précédent. Chercher le nasion en remontant avec l'ongle le dos du nez jusqu'au léger ressaut que fait le bord inférieur du frontal" ((20), p. 570). The anatomical points for the craniological measurements are defined, "en haut: nasion; en bas: bord inférieur de la mandibule, dans le plan médian" ((20), p. 564, No. 11).

Hrdlička calls it the "chin-nasion height" in his work on the Egyptians ((4), p. 61), and describes it as "the distance from the lowest point of the chin in the median line to a point corresponding to the middle of the fronto-nasal articulation." In his later work on Anthropometry ((5), p. 72) he calls it the "Anatomical Face Length" and defines it as "the distance from the menton (the lowest point in the middle of the bony chin) to the nasion."

It may be identified with Martin's ((14), p. 165) No. 18 "Morphologische Gesichtshöhe (Nasomentale Gesichtshöhe; Kieferhöhe; hauteur naso-mentonnière; total face length): Geradlinige Entfernung des Nasion vom Gnathion" and with Schmidt's ((25), p. 107) No. 43 "der Abstand des Kinnes von der Nasenwurzel."

Koganei calls it "Gesichtshöhe B" after Virchow, and describes it as the chin-nasion distance: "die Entfernung der Nasenwurzel (Stirn-Nasennaht) vom unteren Kinnrande" ((8), p. 254). Kubo uses the name "Morphologische Gesichtshöhe" ((10), No. 21) but gives the same definition as Koganei. Von Luschan calls it the "Nasenwurzel-Kinn" distance without giving further details. Sarasin uses the name "Nasomentale Höhe," and as he usually follows Martin, his measurement may be taken to be the same as No. 18 of Martin. Shirokogoroff uses the name "Anatomical Face Length"; I have therefore identified his measurement with this present one which I am calling morphological (or total) face height.

28. It will be noticed that the lower point of reference is spoken of indifferently as the "lower edge of the chin," "the middle of the bony chin," "the gnathion," etc., and while there is general agreement about the region concerned, I have no information to judge whether they all refer in actual practice to the same identical point. As usual the authorities either give no directions at all, or else inadequate directions, regarding the point whether pressure is to be used or not in locating the gnathion and, if so, what degree of pressure. But assuming that there is no ambiguity about the lower point of reference it is clear that the same confusion exists about the upper point as in the case of the nasal height. I believe we may safely assume that each observer will stick to his own particular definition of the "nasion" in taking both measurements (i.e. "the nasal height" and "the morphological (total) face height"). On this assumption and ignoring the important question of pressure we obtain two distinct groups:

(i) Maltese, New Caledonian, Mongolian†, and possibly Korean.

(ii) Cretan, Egyptian, Old American, and possibly Aino and the Swedish data.

* This is of course quite a different measurement from No. 8 "Hauteur totale du visage" which corresponds to the Physiognomic Face Length.

† Shirokogoroff (Mongolian) may now be classed with Buxton and Sarasin.

It will perhaps be moderately safe to compare measurements within the same group, but a comparison between series in different groups would be quite invalid*.

(11) *Bizygomatic Breadth.*

29. B.A. Report ((21), p. 9) calls it No. B 5 "Maximum interzygomatic breadth" and defines it as the "maximum diameter between corresponding points on the opposite zygomatic arches," and notes that "the pressure used is to be as much as can be comfortably borne by the person under examination." Buxton calls it "Bizygomatic Breadth" and Shirokogoroff uses the name "Interzygomatic Breadth."

The Monaco Commission, No. 6 "Diamètre Bizygomatique. Même technique que pour le crâne. Chercher avec soin le maximum qui est souvent plus en arrière qu'on ne le suppose" ((20), p. 570). The craniological definition specifies the anatomical points as "face externe des apophyses zygomatiques" ((20), p. 564, No. 8).

Hrdlička calls it the "maximum bizygomatic diameter" and defines it as the distance between "the most widely separated points on the external surface of the zygomatic arches," which is virtually the same as the definition of the International Commission's No. 6.

This measurement corresponds to Koganei's "Gesichtsbreite *a*" which he defines as "der grösste Abstand der Jochbogen voneinander" ((8), p. 254). Kubo gives an identical definition but uses the name "Jochbogenbreite," which is also the name used by Martin ((14), p. 160), No. 6 "Jochbogenbreite (oft fälschlich Jochbreite genannt, Gesichtsbreite A nach Virchow; largeur totale de la face ou distance bi-zygomatique; maximum interzygomatic breadth): Geradlinige Entfernung der beiden Zygia, d. h. der am meisten seitlich vorstehenden Punkte der Jochbogen voneinander."

Von Luschan and Sarasin both have measurements under the name "Jochbreite." From the above-mentioned remark of Martin that "Jochbogenbreite" is often wrongly called "Jochbreite," and from the fact that both von Luschan and Sarasin use this measurement to calculate the Facial Index, we may conclude that their "Jochbreite" is the same as our "Bizygomatic Breadth."

30. A difficulty remains about the use of pressure in taking measurements. The B.A. Report alone gives instructions on this point, all other authorities remaining silent. In the absence of any information as to pressure we may, perhaps, assume that all the measurements are comparable, but it must not be forgotten that there is a possibility of a statistically significant error being introduced when we make such an assumption.

(12) *Morphological Facial Index.*

31. The Morphological Facial Index is defined as the ratio of the Morphological Face Height to the Bizygomatic Breadth multiplied by 100. The same comparisons

* With a mean S.D. of 6·4 mm. and a difference of 4 mm. (Havelock Charles's value for the difference between the two definitions of the nasion), we get a difference of 22 times the variance in $(M - M')^2$ in samples of 100.

will therefore be possible as in the case of the Morphological Face Height. That is, we have two distinct groups, and this only if we neglect any pressure difference:

(i) Maltese, New Caledonian, Mongolian, and possibly Korean.

(ii) Cretan, Egyptian, Old American, and possibly Aino and the Swedish data.

(13) *Upper Face Length.*

32. B.A. Report ((21), p. 9): No. B 1 "Upper Face Length from the nasion to the edge of the gum between the two upper central incisor teeth. A contact measurement," which may correspond to Martin's ((14), p. 165) No. 20 "Morphologische Obergesichtshöhe (diamètre nasio-alvéolaire, upper face length): Geradlinige Entfernung des Nasion vom Prosthion," and to the International Commission's No. 11 "Diamètre nasio-alvéolaire." Same technique as on crania for which the anatomical points are defined: "en haut: nasion; en bas: le point le plus inférieur du bord alvéolaire, entre les deux incisives médianes et supérieures" ((20), p. 565, No. 12). The "Upper Face Length"* of Buxton (Maltese) and Shirokogoroff (Mongolian) are therefore comparable†.

33. Koganei ((8), p. 254) has one measurement which he calls the "Mittelgesichtshöhe" after Virchow, and defines as "die Entfernung der Nasenwurzel von der Mundspalte (bei geschlossenen Munde)." Kubo ((10), p. 259) gives an identical definition, and his measurements are therefore comparable with Koganei's. This measurement corresponds to Martin's ((14), p. 165) No. 20 "Physiognomische Obergesichtshöhe (Mittelgesichtshöhe nach Virchow, diamètre nasio-buccal): Geradlinige Entfernung des Nasion vom Stomion," and to the International Commission's No. 10 "Diamètre naso-buccal. Points anatomiques dans le plan médian: en haut: nasion; en bas: interligne des lèvres" ((20), p. 570). This measurement is therefore quite different from the "Upper face length" of the B.A. Report. The measurements of Koganei (Aino) and Kubo (Korean) while comparable with each other are not therefore comparable with those of Buxton (Maltese) or Shirokogoroff (Mongolian).

(14) *Minimum Frontal Diameter.*

34. B.A. Report ((21), p. 2): No. A 3 "Minimum Frontal Diameter. From one frontal crest to the other across the narrowest part of the forehead." Buxton, Shirokogoroff and Hrdlička use the same name, and the latter defines it ((5), p. 73) as "the shortest horizontal diameter between the two temporal crests," which is also virtually the same definition as that given for the International Commission's No. 4 "Largeur frontale minima" ((20), p. 570), the technique being the same as

* The use of the word "length" in connection with this measurement is a little inconsistent, for the word "height" is used for an exactly analogous measurement, namely the "morphological (total) face height."

† We have here a good illustration of the confused terminology at present in use. We speak of Morphological Face *Height* and Upper Face *Length*! It would be well to keep the term Height for projected lengths taken in the vertical, i.e. perpendicular to the standard horizontal plane.

on the skull: "C'est le diamètre horizontal le plus court entre les deux crêtes temporales du frontal" ((20), p. 564, No. 5).

It corresponds to Martin's ((14), p. 159) No. 4, "Kleinste Stirnbreite: Geradlinige Entfernung der beiden Frontotemporalia voneinander." Von Luschan uses the same name "Kleinste Stirnbreite," as also Koganei ((8), p. 254, No. 4) and Kubo ((10), p. 240, No. 19), both of whom define it as "der kleinste Abstand der Schläfenlinien am Stirnbein."

I should note here that Martin explicitly states that the instrument should be placed in such a way that no pressure is exerted ("ohne zu drücken"), while all other authorities are silent on this point. I do not know what difference will be made by the application of different degrees of pressure, but if we could neglect such differences then all the measurements would be comparable.

(15) *Bigonial Diameter.*

35. B.A. Report ((21), p. 9): No. B 8 "Gonial Breadth. The diameter between the extreme outer points of the angles of the lower jaw. The measurement is to be taken with the maximum comfortable pressure." Martin's definition ((14), p. 161) is: "No. 8 Unterkieferwinkelbreite (Untergesichtsbreite, untere oder mandibuläre Gesichtsbreite; largeur mandibulaire, bigoniaque ou bigonial; bigonial breadth): Geradlinige Entfernung der beiden Gonia voneinander. Die Tasterspitzen sind nicht hinten an die Unterkieferwinkel, sondern vielmehr etwas seitlich unmittelbar oberhalb des Randes aufzusetzen, damit die äussere seitliche Ausladung der Unterkieferwinkel, die sehr stark sein kann, mitgemessen wird." Nothing is said of the degree of pressure to be used in taking the measurement but, otherwise, it appears to be comparable with the B.A. "gonial breadth" and with the "diamètre bigonique" of the International Commission ((20), p. 570, No. 7).

Buxton and Shirokogoroff call it "Bigonial Breadth," while Hrdlička calls it "Bigonial Diameter" and defines it ((5), p. 73) as the distance between "the gonions or points of the angles of the lower jaw."

Koganei ((8), p. 254, No. 10) calls it "Gesichtsbreite C," and defines it as "der Abstand der Kieferwinkel voneinander." Kubo ((10), p. 279, No. 27) uses the name "Unterkieferwinkelbreite" and describes it as Virchow's "Gesichtsbreite C." Von Luschan calls it "Kieferwinkelbreite."

It is by no means clear that all these measurements were taken between the outer points of the angles as Martin directs and, indeed, all the definitions of the gonia of the living face that have been given are lacking in precision and they might be interpreted in different ways by different observers. Another source of uncertainty is introduced by the absence of instructions regarding the question of pressure and for this measurement that is a vital point. No pair of these bigonial diameters can be compared without making assumptions which may be unjustified.

(16) *Internal Ocular Breadth.*

36. B.A. Report ((21), p. 9): No. B 7 "Internal Ocular Breadth. The diameter between the two internal canthi of the eyelids. A contact measurement." It adds

the direction that the measurement is to be taken when the eyes are open. Buxton who uses the same name, and Shirokogoroff who has one measurement called "Internal Interocular Breadth," have probably followed the above definition which is virtually the same as the International Commission's No. 16 "Largeur bipalpébrale interne. Points anatomiques: angle interne de chaque œil, sans s'occuper de la caroncule" ((20), p. 571). Martin ((14), p. 161) gives it as No. 9 "Breite zwischen den inneren Augenwinkeln (Nasenwurzelbreite, obere Nasenbreite; largeur interoculaire, bi-caronculaire, bi-angulaire interne, bi-oculaire interne ou bi-palpébrale; internal ocular or interocular breadth): Geradlinige Entfernung der beiden Entokanthia voneinander bei offener Lidspalte." Koganei ((8), p. 255, No. 11) describes it as "Distanz zwischen den inneren Augenwinkel" which is probably the same as the above described measurement.

37. Kubo on the other hand is a little uncertain. He calls it ((10), p. 295, No. 30) "Breite zwischen den inneren Augenwinkeln (Interocularabstand)" which would apparently identify his measurement with the above-mentioned sets. But in his description he distinctly mentions the "Abstand der inneren Orbitalränder des knöchernen Schädels" which is taken verbatim from Schmidt ((25), p. 106, No. 38). This would imply that he was actually measuring the distance between the two internal edges of the orbital ridge, which distance of course may differ a little from the distance between the two internal canthi. Comparison with other sets is therefore a little doubtful.

(17) *External Ocular Breadth.*

38. B.A. Report ((21), p. 9): No. B 6 "External Ocular Breadth. The diameter between the two external canthi of the eyelids. A contact measurement." Measurement to be taken when the eyes are open.

Shirokogoroff who calls it ((27), p. 1, No. 19) "External interocular breadth" has probably followed the above definition. Koganei describes it ((8), p. 255, No. 12) as "Distanz der äusseren Augenwinkeln," and his measurements are probably comparable with those of Shirokogoroff.

The above measurement corresponds to Schmidt's ((25), p. 106) No. 37 "Breite zwischen den beiden äusseren Augenwinkeln," to the International Commission's No. 15 "Largeur bi-palpébrale externe. Points anatomiques: angle externe de chaque œil, dans sa région profonde, en contact immédiat avec le globe de l'œil. Les yeux du sujet étant bien ouverts, le regard un peu au-dessus de l'horizon, viser ce point avec les branches du compas appuyées sur les joues du sujet" ((20), p. 571), and to Martin's ((14), p. 161) No. 10 "Breite zwischen den äusseren Augenwinkeln (Obergesichtsbreite nach Weisbach; largeur bi-oculaire externe, bi-angulaire, bi-palpébrale externe; external biocular breadth): Geradlinige Entfernung der beiden Ektokanthia voneinander bei offener Lidspalte."

39. Kubo is again doubtful. He calls it ((10), p. 301, No. 31) "Breite zwischen den äusseren Augenwinkeln," but actually defines it as "Breite zwischen den

äusseren knöchernen Augenhöhlenrändern." This of course is a quite different measurement corresponding to the B.A. Report ((21), p. 9) No. B 5 "External Orbital Breadth. Maximum diameter between the outer margins of the orbital openings," to Schmidt's ((25), p. 106) No. 36 "Breite zwischen den äusseren knöchernen Augenhöhlenrändern," and to Martin's ((14), p. 162) No. 10 (1) "Breite zwischen den äusseren Augenhöhlenrändern (diamètre orbitaire; external orbital breadth): Geradlinige Entfernung der beiden äusseren Orbitalränder im Niveau der äusseren Lidkommissar voneinander." Kubo's measurement* cannot therefore be compared with either Shirokogoroff's or Koganei's, but it may be compared with Buxton's who definitely took the "External Orbital Breadth" (B 5 of the B.A. Report) and not the "External Ocular Breadth" (B 6).

(18) *Breadth of the Mouth.*

40. Hrdlička defines it as the "distance between the angles of the mouth at points where the mucous membrane joins the skin, with mouth naturally closed, without tension" ((5), p. 75), which is practically a literal translation of the International Commission's No. 17 "Largeur de la bouche. Points anatomiques: commissures des lèvres, au point où la muqueuse se continue avec la peau; prendre leur distance, la bouche étant dans sa position moyenne" ((20), p. 571).

Koganei ((8), p. 255, No. 16) calls it "Länge des Mundes" but does not give any definition, Kubo calls it ((10), p. 339, No. 37) "Breite (Länge) des Mundes," and notes that it was measured with the mouth closed. The addition of the word "Länge" within brackets shows that it was formerly known by that name. It corresponds to Martin's ((14), p. 163) No. 14 "Breite der Mundspalte (Mundlänge; longueur buccale): Geradlinige Entfernung der beiden Cheilia voneinander. Der Mund muss geschlossen und in Ruhelage sein." All the measurements are therefore comparable.

(19) *Length of the Ear.*

41. B.A. Report ((21), p. 10): D 2 "Greatest length of the ear. From the highest to the lowest point of the auricle." Shirokogoroff uses the same name.

Monaco Commission: No. 19 (*a*) "Oreille—Longueur maxima: Points anatomiques: en haut: le point le plus élevé du bord de l'hélix; en bas: extrémité inférieure du lobule. Placer la tige du compas parallèle au grand axe de l'oreille et ses branches tangentes aux points indiqués, sans presser" ((20), p. 572).

Hrdlička called it "Height of the Ear†" in his paper on the Egyptians ((4), p. 91), and stated that his measurement was the same as that of Topinard ((29), 1885, p. 1004), Weisbach (*Zeit. für Ethnologie*, Bd. IX, supplement, 1878), and

* This is another example of a measurement being called by a name properly belonging to an altogether different measurement.

† It would be well to retain this term for the projection of the maximum length on the vertical, i.e. when the rod of the callipers is *held perpendicular*, with the head adjusted to the standard horizontal plane.

Schwalbe (Beiträge zur Anthropologie des Ohres—Virchow's *Festschrift*, 1891, p. 95). Schwalbe himself called it "Grösste Länge des Ohres" ((26), pp. 113—114) and mentioned that "dieses Mass ist identisch mit dem, welches von Weisbach und Virchow als Höhe des Ohres bezeichnet wird. Die grösste Länge des Ohrscheitels bis zum entferntesten Punkt des unteren Ohrrandes." Comparing Schwalbe's diagram ((26), p. 115, Fig. 10, $ab = L$) with that of the B.A. Report (fig. 1, AB), I find that the two definitions are identical.

42. In his later book ((5), p. 76) Hrdlička changed the name to "Maximum Length" and defined the landmarks as follows: "Superiorly the highest point on the border of the helix; inferiorly the lowest point on the lobule. The rod of the compass should be held parallel to the long axis of the ear; use no pressure." This is again practically a translation of the International Commission's No. 19 (*a*).

43. Koganei calls it ((8), p. 255, No. 17) "Länge des Ohres," and defines it: "vom unteren Ende des Ohrläppchens bis zum hochsten Punkte der Ohrmuschel gemessen."

Kubo calls it ((10), p. 349, No. 39) "Physiognomische Länge des Ohres (Höhe des Ohres)" and following von Luschan ((11), p. 40, No. 54) defines it as "dessen wirkliche Höhe zwischen den wagerechten Tangenten an dem oberen und dem unteren Rand." The measurement corresponds to Martin's ((14), p. 168) No. 29 "Physiognomische Länge des Ohres (grösste Länge des ganzen Ohres nach Schwalbe): Geradlinige Entfernung des Ohrscheitels (Superaurale) von dem tiefsten Punkte des Ohrläppchens (Subaurale)." Probably all the measurements are comparable.

III. The Comparative Value of the Material.

44. It may be presumed that the chief, if not the only, reason for taking anthropometric measurements is to use such measurements for purposes of comparison, that is: for the study of racial or group resemblances and divergences. The real value of any particular collection of material would then be determined by the possibilities of comparison offered by it. Any portion of the material which is not available for comparison may therefore be considered useless for our purposes.

45. Let us consider the present material. Counting Shirokogoroff's material as one single unit we have nine distinct sets of observations taken by nine different observers. For any particular character for which comparable measurements are available for every set, $\frac{1}{2}(9 \times 8) = 36$ different comparisons will be theoretically possible. Owing to lack of agreement in definitions or technique all measurements however may not be comparable, in which case the number of comparisons actually possible will be considerably reduced. For example, for the Nasal Index we have the following groups:

(i) Korean (Kubo) (ii) Aino (Koganei) (iii) Mongolian (Shirokogoroff) (iv) Swedish (Lundborg)	Single sets, no comparison possible.

(v) Maltese (Buxton)
New Caledonian (Sarasin) } Two sets, 1 comparison possible.

(vi) Cretan (von Luschan)
Egyptian (Hrdlička)
Old American (Hrdlička) } Three sets, 3 comparisons possible.

Thus only 4 (out of a possible 36) comparisons are feasible in practice. We may reckon the effective value of the material to be 4/36 or about 11 °/₀ only.

I give below, Table I, such detailed analysis for each character separately. Col. (3) gives the total number of sets for which a particular measurement is available, Col. (4) the number of comparisons which would have been possible if all measurements had been strictly comparable, Col. (5) the number of comparisons which might be reasonably made in practice, though even then they include several—such as those between the bigonial diameters (see Section (15) above)—which could only be justified fully if far more adequate information were available

TABLE I.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
*1	Head Length	9	36	9	25 °/₀	12	33 °/₀
*2	Head Breadth	9	36	9	25	16	44
*3	Cephalic Index	9	36	9	25	12	33
4	Height of Head	5	10	2	20	2	20
5	Horizontal Circumference of Head	2	1	1	100	1	100
*6	Nasal Height	9	36	4	11	13	36
*7	Nasal Breadth	9	36	16	44	22	61
*8	Nasal Index	9	36	4	11	4	11
9	Nasal Depth	2	1	1	100	1	100
*10	Morphological Face Height ...	9	36	6	17	16	44
*11	Bizygomatic Diameter	9	36	36	100	36	100
*12	Morphological Facial Index ...	9	36	6	17	16	44
*13	Upper Face Length	4	6	2	33	2	33
14	Minimum Frontal Diameter ...	8	28	28	100	28	100
15	Bigonial Diameter	7	21	21	100	21	100
16	Internal Ocular Breadth	5	10	6	60	6	60
*17	External Ocular Breadth ...	4	6	2	33	2	33
18	Breadth of Mouth	5	10	10	100	10	100
19	Length of Ear	5	10	10	100	10	100
	Total		427	182	42·6 °/₀	230	53·8 °/₀

regarding the definitions of points and measurements. Lacking any direct negative evidence, it has been assumed that the different workers used the same degree of pressure in determining a particular measurement. If it were demanded that the comparison should be perfectly rigorous, so that the absence of information on such a point would necessitate the rejection of the observation, then the strictly comparable material would be reduced almost to none. Col. (6) is the percentage effective value (or efficiency) given by the ratio of (5) to (4) expressed as a

percentage, Col. (7) and Col. (8) corresponding numbers if all doubtful cases are included as being available for comparison.

Adding up for all characters we notice that only 182 comparisons can be made with any degree of certainty* out of a possible 427, giving an effective value of 42·6°/₀. If doubtful cases were included the actual number and percentage would be increased to 230 and 53·8°/₀ respectively. The above analysis does not however give a fair idea of the real position, for in the above list we have included a number of characters which are rather rare.

46. Let us confine ourselves to the standard list given in *Notes and Queries* (19). It consists of 9 measurements of the head, out of which we have 8 in the above list. Including the three relevant indices we obtain a total of 11 measurements (marked with an asterisk in Table I) for which we now get Table II.

TABLE II.

Number of Comparisons

Theoretically possible	Actually possible	
	Rejecting doubtful cases	Including doubtful cases
336	103	151
Percentage	30·6°/₀	45·0°/₀

If we further omit the two characters Nos. (13) and (17) in Table I, i.e. restrict ourselves to the *characters which are available in every set*, we get the following table.

TABLE III.

Number of Comparisons

Theoretically possible	Actually possible	
	Rejecting doubtful cases	Including doubtful cases
324	99	147
Percentage	30·6°/₀	45·4°/₀

The appalling fact then comes out that only from 30°/₀ to 45°/₀ of the theoretically possible comparisons can actually be made in practice, i.e. fully from

* Even this assumes that the same amount of pressure has been used by different observers in measuring Morphological Face Height, Bizygomatic Diameter, Minimum Frontal Diameter and Bigonial Diameter. If different degrees of pressure were employed, and such differences cannot be legitimately ignored, then the actual number of possible comparisons is reduced to 85 or to less than 20°/₀ of the number theoretically possible. Not neglecting pressure differences the possible comparisons in Table II are 55 or 16·4°/₀ and in Table III, 51 or 15·7°/₀!

55°/₀ to 70°/₀ of the usefulness of the material is wasted owing to lack of agreement in definitions and technique. A more exacting critic might be unwilling to make some of the assumptions which we have made in order to obtain these figures. If he demand that comparison should be restricted to measurements *certainly known* to have been found by identical methods then he would find that only 20°/₀ or even 15°/₀ of the possible comparisons could be made legitimately. To put it in a different way *the effective value of the present material might have been increased at least two- or three-fold and possibly five- or six-fold by proper standardisation without any additional expenditure of energy, time or money in collecting the material.*

47. I may note here that the present material was originally collected with a view to obtaining as large a number of comparable characters as possible. In fact I rejected several series of observations where a preliminary examination showed that a sufficient number of comparable characters would not be available. Thus my very generous estimate of a wastage of from 55°/₀ to 70°/₀ is more likely to be an under-estimate than otherwise for measurements in general.

The actual situation is indeed much worse owing to another reason. Ordinarily one gets from 10 to 15 different measurements (on the head) in good sets of observations, and only rarely 20 or more. So that owing to the reduction in number of measurements due to lack of standardisation the number of characters actually available for comparison may be easily reduced to anything from 5 to 8 on an average. But for serious comparative studies it is absolutely essential to have a larger number of characters, preferably 20 or more. Thus the reduction in number of available characters, owing to want of standardisation, will very often render the whole material practically useless for comparative purposes.

We are obliged to conclude that even from the best of the series of head measurements at present available we are quite unable to establish reliable inter-racial comparisons. The labour of estimating how far the measurements may be legitimately compared is not one from which the biometrician is likely to derive much profit, for he will usually be left with insufficient material to serve his needs. The comparisons frequently made between similar measurements taken according to different methods of technique are merely idle.

IV. Conclusions.

48. The present study indicates that the greater part of the existing anthropometric measurements on the living are probably useless for comparative purposes owing to lack of agreement in definitions and technique.

Such lack of standardisation is due to the existence of a multiplicity of "standard lists" each with its own adherents. The International Agreement of 1906 has failed to achieve any appreciable degree of uniformity among anthropometric workers. Out of eight observers considered in this paper (all of whom carried out their present investigations after 1906) only two (Hrdlička and Shirokogoroff) mention the 1906 Agreement, but both of them introduce departures from this Agreement. Several

important standard lists, e.g. the British Association Report (1908), *Notes and Queries* (1912), Martin's list (1914), were drawn up after the International list, but have made substantial departures from it.

The failure of the International Agreement of 1906 may be ascribed to several causes. Specifications are in many places vague and uncertain, detailed descriptions are not often given about the actual procedure to be followed in taking measurements, and the list contains too few characters. Finally it does not appear to have been readily available in a convenient form.

It is hardly necessary to point out that no progress in comparative Anthropometry will be possible until arrangements have been made for securing adequate uniformity of standards among different observers. The only way in which this can be achieved is to convene another International Conference for drawing up a more satisfactory set of standards than the Monaco (1906) Agreement.

The greater part of this paper was written in the Biometric Laboratory, University of London, and I have to thank Prof. Karl Pearson for his active interest and valuable suggestions at every stage of my work. I also acknowledge with gratitude the help I have received from Dr G. M. Morant of the same laboratory.

APPENDIX I.

The following table gives a rough idea of the order of the difference made in $Z = (M - M')/\Sigma_{(M-M')}$ for certain estimated values of the difference in means arising from differences in technique. Col. (3) gives the total number of individual measurements (taken from the data considered in the present paper) on which the average

(1) No.	(2) Name of character	(3) No. of Individuals	(4) $\bar{\sigma}$ Mean Standard deviation	(5) Estimated difference in Means	(6) "Z" for samples of 100
1	Head Length	1863	6·92 mm.	3 mm.	3·0
2	Head Breadth	1864	5·72 mm.	3 mm.	3·7
3	Cephalic Index	1863	4·20	0·4	0·7
4	Height of Head	795	6·13 mm.	3 mm.	3·4
5	Horizontal Circumference of Head	640	13·25 mm.		
6	Nasal Height	1856	3·74 mm.	4 mm.	7·5
7	Nasal Breadth	1852	2·93 mm.	5 mm.	12·1
8	Nasal Index	1850	7·97	5·0	4·4
9	Nasal Depth	440	1·95 mm.		
10	Morphological Face Height ...	1852	6·41 mm.	4 mm.	4·4
11	Bizygomatic Diameter	1855	5·45 mm.		
12	Morphological Facial Index ...	1847	4·72	3·0	5·7
13	Upper Face Length	592	3·81 mm.	5 mm.	9·2
14	Minimum Frontal Diameter ...	1612	4·96 mm.		
15	Bigonial Diameter	1620	6·07 mm.		
16	Internal Ocular Breadth	1510	2·91 mm.	3 mm.	7·2
17	External Ocular Breadth ...	1185	4·87 mm.	3 mm.	4·3
18	Breadth of Mouth	866	3·72		
19	Length of Ear	735	4·66 mm.		

standard deviation ($\bar{\sigma}$) given in Col. (4) is based. Col. (5) gives the estimated difference in means due to differences in technique. I may note that these are on the whole conservative estimates. Col. (6) gives "*Z*" for samples of 100. It will be seen that except for the Cephalic Index differences in technique will certainly not be negligible in comparison with errors of sampling when the group consists of 100 individuals or more.

Appendix II.

Bibliographical References.

1. Broca, Paul. *Instructions générales pour les recherches anthropologiques à faire sur les vivants.* Paris, 1879.
2. Charles, R. Havelock. Nasal Index Compared upon the Head and Skull, etc. *Journal Asiatic Society of Bengal*, Vol. lxiii, Part III, No. 1. Calcutta, 1894.
3. Dudley-Buxton, L. H. Ethnology of Malta and Gozo. *Journ. Royal Anthrop. Inst. London*, Vol. lii, 1922.
4. Hrdlička, Aleš. The Natives of Kharga Oasis, Egypt. *Smithsonian Institute Miscellaneous Collections*, Vol. lix, No. 1. Washington, 1912.
5. —— *Anthropometry.* Wistar Institute of Anatomy and Biology. Philadelphia, 1920.
6. —— *Old Americans.* Williams Wilkins. Baltimore, U.S.A., 1925.
7. International Agreement for the Unification of Anthropometric Measurements to be made on the Living Subject. *Report of the Commission appointed by the XIVth International Congress of Pre-historic Anthropology and Archaeology at Geneva* (1912). English translation by W. L. H. Duckworth. Cambridge, 1912.
8. Koganei, Y. Beiträge zur physischen Anthropologie der Aino. *Mitt. d. med. Fakult. d. kais. Univ. Tokio*, Bd. ii. 1893.
9. —— Messungen an Chinesischen Soldaten. *Mitt. d. med. Fakult. d. kais. Univ. Tokio*, Bd. vi. 1902.
10. Kubo, T. Beiträge zur physischen Anthropologie der Koreaner. *Mitt. d. med. Fakult. d. kais. Univ. Tokio*, Bd. xii. 1913.
11. Luschan, F. von. Neumayer's *Anleitung zu wissen. Beobachtungen auf Reisen*, Bd. ii. 1906.
12. —— Beiträge zur Anthropologie von Kreta. *Zeit. für Ethnologie.* Berlin, 1913.
13. Lundborg, H. and Linders, F. J. *Anthropologica Suecica.* Swedish State Institute for Race Biology, Upsala, 1926.
14. Martin, Rudolf. *Lehrbuch der Anthropologie.* Jena, 1914.
15. —— *Die Inlandstämme der malayischen Halbinsel.* Jena, 1905.
16. *Notes and Queries on Anthropology*, etc. 1st edition, edited by Col. Lane-Fox and Lieut.-Gen. Pitt-Rivers. Anthrop. Institute, London, 1874.
17. —— 2nd edition, edited by J. G. Garson and C. H. Read. Anthrop. Institute, London, 1892.
18. —— 3rd edition, edited by J. G. Garson and C. H. Read. Anthrop. Institute, London, 1899.
19. —— 4th edition, edited by B. Freire-Marreco and J. L. Myers. Anthrop. Institute, London, 1912.

20. "Report of the Anthropometric Commission appointed by the XIIIth International Congress of Pre-historic Anthropology and Archaeology. Monaco, 1906." *L'Anthropologie*, Vol. XVII, 1906, pp. 559—572; also English translation in Hrdlička's *Anthropometry*, pp. 10—24.

21. *Report of the Anthropometric Committee of the British Association.* British Association for the Advancement of Science. Dublin, 1908.

22. SARASIN, FRITZ. *Anthropologie der Neu-Caledonier und Loyalty-Insulaner.* Berlin, 1916—1922.

23. —— *Versuch einer Anthropologie der Insel Celebes.* Wiesbaden, 1906.

24. *Schedule of Measurements in Anthropology.* British Association for the Advancement of Science. 1895.

25. SCHMIDT, E. *Anthropologische Methoden.* Leipzig, 1888.

26. SCHWALBE, G. Beiträge zur Anthropologie des Ohres. Virchow's *Festschrift*, Bd. I. 1891.

27. SHIROKOGOROFF, S. M. *Anthropology of Northern Asia.* Royal Asiatic Society, North China Branch. Extra Vol. II. Shanghai, 1923.

28. —— *Anthropology of Eastern China, etc.* Royal Asiatic Society, North China Branch. Extra Vol. IV. Shanghai, 1925.

29. TOPINARD, PAUL. *Éléments d'Anthropologie générales.* Paris, 1885.

30. —— *Anthropology.* English translation. London, 1890.

31. VIRCHOW, R. List in *Zeit. für Ethnologie*, 1885. Verhandlungen der Berliner Gesellschaft für Anthropologie, Ethnologie und Urgeschichte, pp. 99—102.

32. EICKSTEDT, E. VON. Beiträge zur Rassenmorphologie der Weichteilnase. *Zeit. für Morph. u. Anthrop.* Bd. XXV, 1926.

THE GENERALISED PRODUCT MOMENT DISTRIBUTION IN SAMPLES FROM A NORMAL MULTIVARIATE POPULATION.

By JOHN WISHART, M.A., B.Sc. Statistical Department, Rothamsted Experimental Station.

1. *Introduction.*

For some years prior to 1915, various writers struggled with the problems that arise when samples are taken from uni-variate and bi-variate populations, assumed in most cases for simplicity to be normal. Thus "Student," in 1908*, by considering the first four moments, was led by K. Pearson's methods to infer the distribution of standard deviations, in samples from a normal population. His results, for comparison with others to be deduced later, will be stated in the form

$$dp = \frac{1}{\Gamma\left(\frac{N-1}{2}\right)} A^{\frac{N-1}{2}} . e^{-Aa} . a^{\frac{N-3}{2}} da \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1),$$

where N is the size of the sample, and

$$A = \frac{N}{2\sigma^2}, \quad a = s^2,$$

σ being the standard deviation of the sampled population, and s that estimated from the sample. Thus, if $x_1, x_2, \ldots x_N$ are the sample values,

$$N\bar{x} = \sum_1^N (x),$$

and

$$Ns^2 = \sum_1^N (x - \bar{x})^2.$$

When bi-variate populations were considered, other problems arose, such as the distribution of the correlation coefficient and of the regression coefficient in samples. These problems, taken by themselves, were found to be difficult, and only approximative results had been reached, when, in 1915, R. A. Fisher† gave a formula for the simultaneous distribution of the three quadratic statistical derivatives, namely the two variances (squared standard deviations) and the product moment coefficient. Thus, let $x_1, x_2, \ldots x_N$ represent the sample values of the x-variate, and $y_1, y_2, \ldots y$ the corresponding values for the y-variate, let σ_1 and σ_2 be the standard deviations of the sampled population and ρ the correlation between x and y. We then calculate the following statistical derivatives from the sample:

$$N\bar{x} = \sum_1^N (x) \qquad N\bar{y} = \sum_1^N (y)$$

$$Ns_1^2 = \sum_1^N (x - \bar{x})^2 \qquad Ns_2^2 = \sum_1^N (y - \bar{y})^2$$

$$Nrs_1s_2 = \sum_1^N (x - \bar{x})(y - \bar{y}).$$

* *Biometrika*, Vol. VI. 1908, pp. 4—6.

† *Biometrika*, Vol. X. 1915, p. 510.

If we put $A=\frac{N}{2\sigma_1^2(1-\rho^2)}$, $B=\frac{N}{2\sigma_2^2(1-\rho^2)}$, $H=-\frac{N\rho}{2\sigma_1\sigma_2(1-\rho^2)}$,

$a=s_1^2$, $b=s_2^2$, $h=s_1s_2r$,

then Fisher's result, for the simultaneous distribution of a, b and h, may be put in the symmetrical form*

$$dp=\frac{1}{\sqrt{\pi}\,\Gamma\left(\frac{N-1}{2}\right)\Gamma\left(\frac{N-2}{2}\right)}\begin{vmatrix} A & H \\ H & B \end{vmatrix}^{\frac{N-1}{2}}.e^{-Aa-Bb-2Hh}.\begin{vmatrix} a & h \\ h & b \end{vmatrix}^{\frac{N-4}{2}}dadbdh \qquad \text{.....(2).}$$

The distribution of the correlation coefficient was deduced by direct integration from this result. Further, K. Pearson and V. Romanovsky, starting from this fundamental formula, were able to deal with the regression coefficients. Pearson, in 1925†, gave the mean value and standard deviation of the regression coefficient, while Romanovsky and Pearson‡, in the following year, published the actual distribution.

New problems arise when we are dealing with three or more variates. There are, for example, the distributions of the partial and multiple correlations, and of the partial regression coefficients. In this domain our knowledge is still far from complete. The partial correlation coefficient has been shown to be distributed exactly as a total coefficient, when the size of the sample has been reduced by the number of variates eliminated§. The actual distribution of the multiple correlation coefficient, for the particular case where no real correlation exists in the sampled population, was shown in 1923‖ to be of the form

$$dp=\frac{\Gamma\left(\frac{N-1}{2}\right)}{\Gamma\left(\frac{N-n}{2}\right)\Gamma\left(\frac{n-1}{2}\right)}(R^2)^{\frac{n-3}{2}}(1-R^2)^{\frac{N-n-2}{2}}d(R^2) \qquad \text{.....(3),}$$

where n is the total number of variates. A recent experimental investigation by the present writer¶ may be regarded as confirming this result, as well as was to be expected on the basis of the inadequate numbers used. On the other hand, L. Isserlis**, in 1917, found approximations to the mean value and standard deviation of the multiple correlation coefficient, in the case of three variates, while recently P. Hall††, for an n-fold system, determined (a) the mean value, and (b) the standard deviation, of the multiple and partial correlation coefficients, as far as terms of the order of $\frac{1}{N}$.

* See Romanovsky, *Comptes Rendus*, Tome 180, 1925, p. 1897.

† *Biometrika*, Vol. XVII. 1925, pp. 195—196.

‡ *Bulletin de l'Académie des Sciences de l'U.R.S.S.*, 1926, p. 646. *Proc. Roy. Soc.* A. 112, 1926, p. 1.

§ R. A. Fisher, *Metron*, Vol. III. Nos. 3—4, 1924, pp. 1—2.

‖ R. A. Fisher, *Phil. Trans.* B. Vol. 213, 1923, p. 91.

¶ *Memoirs of the Royal Meteorological Society*, Vol. II. pp. 29—37, 1928. *Rothamsted Memoirs*, Vol. XIV.

** *Phil. Mag.* Vol. XXXIV. 1917, pp. 205—220.

†† *Biometrika*, Vol. XIX. 1927, pp. 100—109.

What is now asserted is that all such problems depend, in the first instance, on the determination of a fundamental frequency distribution, which will be a generalisation of equation (2). It will, in fact, be the simultaneous distribution in samples of the n variances (squared standard deviations) and the $\frac{n(n-1)}{2}$ product moment coefficients. It is the purpose of the present paper to give this generalised distribution, and to calculate its moments up to the fourth order. The case of three variates will first be considered in detail, and thereafter a proof for the general n-fold system will be given.

2. *Tri-variate Product Moment Distribution.*

Let the frequency distribution of the population sampled be

$$dp = \frac{1}{(2\pi)^{\frac{3}{2}}\sigma_1\sigma_2\sigma_3\sqrt{\Delta}} \times e^{-\frac{1}{2\Delta}\left[\frac{(x-m_1)^2}{\sigma_1^2}\Delta_{11}+\frac{(y-m_2)^2}{\sigma_2^2}\Delta_{22}+\frac{(z-m_3)^2}{\sigma_3^2}\Delta_{33}+2\frac{(y-m_2)(z-m_3)}{\sigma_2\sigma_3}\Delta_{23}+2\frac{(z-m_3)(x-m_1)}{\sigma_3\sigma_1}\Delta_{31}+2\frac{(x-m_1)(y-m_2)}{\sigma_1\sigma_2}\Delta_{12}\right]} dxdydz \quad \ldots\ldots\ldots\ldots(4),$$

where Δ is the determinant $|\rho_{st}|$ $s, t = 1, 2, 3$, and Δ_{st} is the minor of ρ_{st} in Δ.

Now let $x_1, x_2, \ldots x_N$ represent the sample values of the x-variate, and

$$y_1, y_2, \ldots y_N, \quad z_1, z_2, \ldots z_N$$

be the corresponding values for the y- and z-variates. Then the chance that

$$x_1, y_1, z_1, x_2, y_2, z_2, \ldots x_N, y_N, z_N$$

should fall within the elementary ranges $dx_1, dy_1, dz_1, \ldots dx_N, dy_N, dz_N$ is

$$dp = \frac{1}{(2\pi)^{\frac{3N}{2}}(\sigma_1\sigma_2\sigma_3)^N\Delta^{\frac{N}{2}}} \times e^{-\frac{1}{2\Delta}\sum_1^N\left[\frac{(x-m_1)^2}{\sigma_1^2}\Delta_{11}+\frac{(y-m_2)^2}{\sigma_2^2}\Delta_{22}+\frac{(z-m_3)^2}{\sigma_3^2}\Delta_{33}+2\frac{(y-m_2)(z-m_3)}{\sigma_2\sigma_3}\Delta_{23}+2\frac{(z-m_3)(x-m_1)}{\sigma_3\sigma_1}\Delta_{31}+2\frac{(x-m_1)(y-m_2)}{\sigma_1\sigma_2}\Delta_{12}\right]} dx_1dy_1dz_1dx_2dy_2dz_2 \ldots\ldots dx_Ndy_Ndz_N \quad \ldots\ldots\ldots\ldots(5).$$

The following statistical derivatives are now to be calculated from the sample and substituted in (5):

$$N\bar{x} = \sum_1^N(x) \qquad N\bar{y} = \sum_1^N(y) \qquad N\bar{z} = \sum_1^N(z)$$

$$Ns_1^2 = \sum_1^N(x-\bar{x})^2, \qquad Ns_2^2 = \sum_1^N(y-\bar{y})^2 \qquad Ns_3^2 = \sum_1^N(z-\bar{z})^2$$

$$Nr_{23}s_2s_3 = \sum_1^N(y-\bar{y})(z-\bar{z}) \quad Nr_{31}s_3s_1 = \sum_1^N(z-\bar{z})(x-\bar{x}) \quad Nr_{12}s_1s_2 = \sum_1^N(x-\bar{x})(y-\bar{y}).$$

In order to transform the element of volume, we require to extend somewhat the geometrical reasoning employed in the two variate case*. The N values of x may be regarded geometrically as specifying a point P in an N-dimensional space,

* See *Biometrika*, Vol. x. pp. 509—510.

whose co-ordinates are $x_1-\bar{x}$, $x_2-\bar{x}$, ... $x_N-\bar{x}$. Similarly the N values of y and the N values of z specify points Q, R in the same space. When $\bar{x}$ and s_1 are fixed, as when a particular sample is chosen, P is constrained to move so that its perpendicular distances from the line $x_1=x_2=\ldots=x_N$ and from the "plane"

$$x_1+x_2+\ldots+x_N=n\bar{x}$$

remain constant. It must therefore lie on the surface of an $N-1$-dimensional sphere which is everywhere at right angles to the radius vector $x_1=x_2=\ldots=x_N$. The element of volume is then proportional to $(\sqrt{N}s_1)^{N-2}\,ds_1 d\bar{x}$. For the factor of proportionality we require the entire area of the surface of a sphere in $N-1$ dimensions. If of radius r this is, for $N=3, 4, 5, 6, 7$,

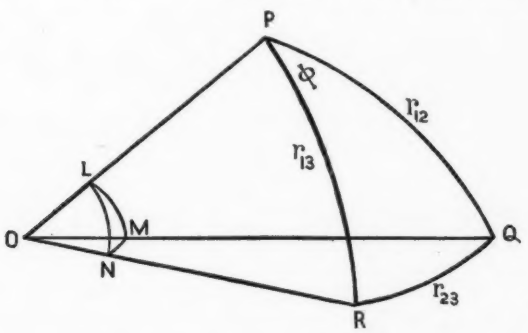
$$2\pi r,\quad 4\pi r^2,\quad 2\pi^2 r^3,\quad \tfrac{8}{3}\pi^2 r^4,\quad \pi^3 r^5,$$

the general result being $\quad 2\,.\,\dfrac{\pi^{\frac{N-1}{2}}}{\Gamma\left(\dfrac{N-1}{2}\right)}\,.\,r^{N-2}.$

Thus we have a contribution to the transformed element of volume of

$$2\,.\,\frac{\pi^{\frac{N-1}{2}}N^{\frac{N-2}{2}}}{\Gamma\left(\dfrac{N-1}{2}\right)}\,.\,s_1{}^{N-2}ds_1 d\bar{x}.$$

By similar reasoning Q and R must lie on concentric spheres in the same space, and there will be corresponding contributions to the transformed element of volume. Let the radii vectores OP, OQ, OR be cut by the unit sphere whose



centre is at O in the points L, M, N. Then LMN is a spherical triangle, specified by the nature of the sample. To find the chance that this particular triangle should be chosen we note that, P being fixed, the chance that Q (or M) should fall within the elementary range $d\theta_1$ (θ_1 being the angle $\dfrac{\pi}{2}-L\hat{O}M$) is equal to

$$\cos^{N-3}\theta_1 d\theta_1 \div \int_0^{\pi}\cos^{N-3}\theta_1 d\theta_1,$$

i.e. to $$\frac{\Gamma\left(\frac{N-1}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{N-2}{2}\right)}\cos^{N-3}\theta_1 d\theta_1.$$

Similarly the chance that R (or N) should fall within $d\theta_2\left(L\hat{O}N=\frac{\pi}{2}-\theta_2\right)$ is equal to

$$\frac{\Gamma\left(\frac{N-1}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{N-2}{2}\right)}\cos^{N-3}\theta_2 d\theta_2.$$

But the points M and N do not vary independently. They are connected by the relation that, if ϕ be the angle between the planes LOM and LON,

$$\cos\phi = r_{23\cdot 1} = \frac{r_{23}-r_{12}r_{13}}{\sqrt{1-r_{12}^2}\sqrt{1-r_{13}^2}}.$$

Now LM being fixed, the chance that LN should fall between the angles ϕ and $\phi+d\phi$, measured from LM, is equal to

$$\sin^{N-4}\phi\, d\phi \div \int_0^\pi \sin^{N-4}\phi\, d\phi = \frac{\Gamma\left(\frac{N-2}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{N-3}{2}\right)}\sin^{N-4}\phi\, d\phi.$$

The transformed volume element will consist of the product of all the above probabilities. The exponential term in (5) is easily expressed in terms of s_1, s_2, s_3 and the r's, and we have

$$dp = \frac{8}{\pi^3\,\Gamma\left(\frac{N-1}{2}\right)\Gamma\left(\frac{N-2}{2}\right)\Gamma\left(\frac{N-3}{2}\right)}\cdot\frac{N^{\frac{3(N-2)}{2}}}{(8\sigma_1^2\sigma_2^2\sigma_3^2\Delta)^{\frac{N}{2}}}$$
$$\times e^{-\frac{N}{2\Delta}\left[\frac{s_1^2+(\bar{x}-m_1)^2}{\sigma_1^2}\Delta_{11}+\frac{s_2^2+(\bar{y}-m_2)^2}{\sigma_2^2}\Delta_{22}+\frac{s_3^2+(\bar{z}-m_3)^2}{\sigma_3^2}\Delta_{33}+2\frac{r_{23}s_2s_3+(\bar{y}-m_2)(\bar{z}-m_3)}{\sigma_2\sigma_3}\Delta_{23}\right.}$$
$$\left.+2\frac{r_{31}s_3s_1+(\bar{z}-m_3)(\bar{x}-m_1)}{\sigma_3\sigma_1}\Delta_{31}+2\frac{r_{12}s_1s_2+(\bar{x}-m_1)(\bar{y}-m_2)}{\sigma_1\sigma_2}\Delta_{12}\right]$$
$$\times s_1^{N-2}s_2^{N-2}s_3^{N-2}\cos^{N-3}\theta_1\cos^{N-3}\theta_2\sin^{N-4}\phi\, d\bar{x}\,d\bar{y}\,d\bar{z}\,ds_1ds_2ds_3d\theta_1d\theta_2d\phi$$
.....(6).

We can integrate out immediately for $\bar{x}$, $\bar{y}$ and $\bar{z}$. The integration produces $(\sqrt{2\pi})^3N^{\frac{3}{2}}\sigma_1\sigma_2\sigma_3\sqrt{\Delta}$, and we get

$$dp = \frac{8}{\pi^{\frac{3}{2}}\,\Gamma\left(\frac{N-1}{2}\right)\Gamma\left(\frac{N-2}{2}\right)\Gamma\left(\frac{N-3}{2}\right)}\left(\frac{N^3}{8\sigma_1^2\sigma_2^2\sigma_3^2\Delta}\right)^{\frac{N-1}{2}}$$
$$\times e^{-\frac{N}{2\Delta}\left[\frac{s_1^2}{\sigma_1^2}\Delta_{11}+\frac{s_2^2}{\sigma_2^2}\Delta_{22}+\frac{s_3^2}{\sigma_3^2}\Delta_{33}+2\frac{r_{23}s_2s_3}{\sigma_2\sigma_3}\Delta_{23}+2\frac{r_{31}s_3s_1}{\sigma_3\sigma_1}\Delta_{31}+2\frac{r_{12}s_1s_2}{\sigma_1\sigma_2}\Delta_{12}\right]}$$
$$\times s_1^{N-2}s_2^{N-2}s_3^{N-2}\cos^{N-3}\theta_1\cos^{N-3}\theta_2\sin^{N-4}\phi\, ds_1ds_2ds_3d\theta_1d\theta_2d\phi$$
.....(7).

Following Romanovsky* in the method of expression, let us put

$$\begin{aligned} s_1^2 &= a \qquad s_1 s_2 \sin\theta_1 &&= h, \\ s_2^2 &= b \qquad s_1 s_3 \sin\theta_2 &&= g, \\ s_3^2 &= c \qquad s_2 s_3 (\cos\theta_1 \cos\theta_2 \cos\phi + \sin\theta_1 \sin\theta_2) &&= f; \end{aligned}$$

$$\left(\text{for } r_{12} = \sin\theta_1,\ r_{13} = \sin\theta_2, \text{ and } r_{23\cdot 1} = \cos\phi = \frac{r_{23} - r_{12} r_{13}}{\sqrt{1 - r_{12}^2}\sqrt{1 - r_{13}^2}}\right).$$

For the Jacobian of this final transformation we have

$$\frac{\partial(a, b, c, f, g, h)}{\partial(s_1, s_2, s_3, \theta_1, \theta_2, \phi)} = 8 s_1 s_2 s_3 \begin{vmatrix} \frac{\partial f}{\partial \theta_1} & 0 & \frac{\partial h}{\partial \theta_1} \\ \frac{\partial f}{\partial \theta_2} & \frac{\partial g}{\partial \theta_2} & 0 \\ \frac{\partial f}{\partial \phi} & 0 & 0 \end{vmatrix}$$

$$= 8 s_1^3 s_2^3 s_3^3 \cos^2\theta_1 \cos^2\theta_2 \sin\phi$$

$$= 8 \{bc(ab - h^2)(ac - g^2) \mid abc\}^{\frac{1}{2}},$$

where $|abc|$ denotes the determinant

$$\begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix}.$$

The element of volume then becomes on transformation

$$\frac{1}{8}(abc)^{\frac{N-2}{2}} \frac{\{(ab - h^2)(ac - g^2)\}^{\frac{N-3}{2}}}{a^{N-3}(bc)^{\frac{N-3}{2}}} \cdot \frac{a^{\frac{N-4}{2}} |abc|^{\frac{N-4}{2}}}{\{(ab - h^2)(ac - g^2)\}^{\frac{N-4}{2}}} \cdot \frac{|abc|^{-\frac{1}{2}}}{(bc)^{\frac{1}{2}} \{(ab - h^2)(ac - g^2)\}^{\frac{1}{2}}}$$

$$= \frac{1}{8} \begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix}^{\frac{N-5}{2}},$$

a very neat result.

Finally we have for the simultaneous distribution of the three variances and the three product moment coefficients the symmetrical expression

$$dp = \frac{1}{\pi^3 \Gamma\left(\frac{N-1}{2}\right) \Gamma\left(\frac{N-2}{2}\right) \Gamma\left(\frac{N-3}{2}\right)} \left(\frac{N^3}{8\sigma_1^2 \sigma_2^2 \sigma_3^2 \Delta}\right)^{\frac{N-1}{2}}$$

$$\times e^{-\frac{N}{2\Delta}\left\{\frac{a}{\sigma_1^2}\Delta_{11} + \frac{b}{\sigma_2^2}\Delta_{22} + \frac{c}{\sigma_3^2}\Delta_{33} + 2\frac{f}{\sigma_2\sigma_3}\Delta_{23} + 2\frac{g}{\sigma_3\sigma_1}\Delta_{31} + 2\frac{h}{\sigma_1\sigma_2}\Delta_{12}\right\}}$$

$$\times \begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix}^{\frac{N-5}{2}} da\,db\,dc\,df\,dg\,dh.$$

* *Comptes Rendus*, loc. cit. See also *Metron*, Vol. v. No. 4, 1925, p. 31.

We may simplify this expression by writing

$$A=\frac{N}{2\sigma_1^2}\cdot\frac{\Delta_{11}}{\Delta},\quad B=\frac{N}{2\sigma_2^2}\cdot\frac{\Delta_{22}}{\Delta},\quad C=\frac{N}{2\sigma_3^2}\cdot\frac{\Delta_{33}}{\Delta},$$

$$F=\frac{N}{2\sigma_2\sigma_3}\cdot\frac{\Delta_{23}}{\Delta},\quad G=\frac{N}{2\sigma_3\sigma_1}\cdot\frac{\Delta_{31}}{\Delta},\quad H=\frac{N}{2\sigma_1\sigma_2}\cdot\frac{\Delta_{12}}{\Delta},$$

when it becomes

$$dp=\frac{1}{\pi^{\frac{3}{2}}\Gamma\left(\frac{N-1}{2}\right)\Gamma\left(\frac{N-2}{2}\right)\Gamma\left(\frac{N-3}{2}\right)}\begin{vmatrix}A & H & G\\ H & B & F\\ G & F & C\end{vmatrix}^{\frac{N-1}{2}}\cdot e^{-Aa-Bb-Cc-2Ff-2Gg-2Hh}\times\begin{vmatrix}a & h & g\\ h & b & f\\ g & f & c\end{vmatrix}^{\frac{N-5}{2}}dadbdcdfdgdh \quad\ldots\ldots\ldots\ldots(8).$$

It is to be noted that $|abc|$ is equal to $s_1^2s_2^2s_3^2\,|r_{pq}|$, $p, q=1, 2, 3$.

This is the fundamental frequency distribution for the three variate case, and in a later section the calculation of its moment coefficients will be dealt with.

3. *Multi-variate Distribution. Use of Quadratic co-ordinates.*

A comparison of equation (8) with the corresponding results (1) and (2) for uni-variate and bi-variate sampling, respectively, indicates the form the general result may be expected to take. In fact, we have for the simultaneous distribution in random samples of the n variances (squared standard deviations) and the $\frac{n(n-1)}{2}$ product moment coefficients the following expression:

$$dp=\frac{\begin{vmatrix}A_{11} & A_{12}\ldots A_{1n}\\ A_{21} & A_{22}\ldots A_{2n}\\ \vdots & \vdots\quad\vdots\\ A_{n1} & A_{n2}\ldots A_{nn}\end{vmatrix}^{\frac{N-1}{2}}}{(\sqrt{\pi})^{\frac{1}{2}n(n-1)}\,\Gamma\left(\frac{N-1}{2}\right)\Gamma\left(\frac{N-2}{2}\right)\ldots\Gamma\left(\frac{N-n}{2}\right)}\times e^{-A_{11}a_{11}-A_{22}a_{22}-\ldots-A_{nn}a_{nn}-2A_{12}a_{12}-2A_{13}a_{13}-\ldots-2A_{n-1\,n}a_{n-1\,n}}\times\begin{vmatrix}a_{11} & a_{12}\ldots a_{1n}\\ a_{21} & a_{22}\ldots a_{2n}\\ \vdots & \vdots\quad\vdots\\ a_{n1} & a_{n2}\ldots a_{nn}\end{vmatrix}^{\frac{N-n-2}{2}}da_{11}da_{12}\ldots da_{nn}\quad\ldots\ldots\ldots\ldots\ldots(9),$$

where $a_{pq}=s_ps_qr_{pq}$, and $A_{pq}=\frac{N}{2\sigma_p\sigma_q}\cdot\frac{\Delta_{pq}}{\Delta}$, Δ being the determinant

$$|\rho_{pq}|,\ p, q=1, 2, 3, \ldots n,$$

and Δ_{pq} the minor of ρ_{pq} in Δ.

This result can be proved by the aid of the following general geometrical considerations. We shall begin by defining the *quadratic co-ordinates* of a set of points, and thereafter develop the argument by the use of these co-ordinates.

(*a*) If $x_{p1}, x_{p2}, \ldots x_{pn}$ are the rectangular co-ordinates of a point (p) in space of n dimensions, then the configuration of a set of n points relative to the origin may be specified by the co-ordinates

$$\xi_{pq} = \sum_{k=1}^{n} (x_{pk} x_{qk}).$$

These co-ordinates will be unchanged by any rotation of the whole system about the origin, the configuration of the set of $(n+1)$ points including the origin being unchanged.

(*b*) The determinant $|\xi_{pq}|$, $p, q = 1, 2, 3, \ldots n$, on expanding the several terms, may be recognised as the square of the determinant $|x_{pq}|$, $p, q = 1, 2, 3, \ldots n$, and is therefore equal to the square of the volume enclosed by completing the parallel-faced figure having one corner at the origin and edges running to the n points of the figure. This volume will be represented by v_n.

(*c*) The perpendicular distance of the point (n) from the plane space passing through the origin and the points (1) to $(n-1)$ is $\dfrac{v_n}{v_{n-1}}$.

(*d*) If the points (1) to $(n-1)$ are fixed, and that of the point (n) is specified by the co-ordinates $\xi_{n1}, \xi_{n2}, \ldots \xi_{nn}$, then the element of volume in the neighbourhood of the point (n) may be represented by

$$\frac{1}{2v_n} d\xi_{n1}, d\xi_{n2} \ldots d\xi_{nn}.$$

For the Jacobian

$$\frac{\partial(\xi_{n1}, \xi_{n2}, \ldots \xi_{nn})}{\partial(x_{n1}, x_{n2}, \ldots x_{nn})} = \begin{vmatrix} x_{11} & x_{12} & \ldots & x_{1n} \\ x_{12} & x_{22} & \ldots & x_{2n} \\ \vdots & \vdots & & \vdots \\ 2x_{1n} & 2x_{2n} & & 2x_{nn} \end{vmatrix} = 2v_n.$$

(*e*) If in a space of N dimensions the points (1) to $(n-1)$ are fixed, and that of the point (n) is specified by the co-ordinates $\xi_{n1}, \xi_{n2}, \ldots \xi_{nn}$ within ranges

$$d\xi_{n1}, d\xi_{n2}, \ldots d\xi_{nn},$$

then the point (n) is free to move on the surface of a sphere in $N-n+1$ dimensions, of which the centre is the foot of the perpendicular from (n) upon the space containing the points (1) to $(n-1)$. The radius of this sphere is therefore $\dfrac{v_n}{v_{n-1}}$, and the area of its surface is

$$2 \cdot \frac{\pi^{\frac{1}{2}(N-n+1)}}{\Gamma\left(\dfrac{N-n+1}{2}\right)} \left(\frac{v_n}{v_{n-1}}\right)^{N-n}$$ (see above, p. 35).

This surface is everywhere normal to the space passing through the origin and the points (1) to (n). But the element of volume in this space is

$$\frac{1}{2v_n}\, d\xi_{n1},\, d\xi_{n2},\, \dots\, d\xi_{nn}.$$

Hence the entire volume that may be occupied by (n) is

$$\frac{\pi^{\frac{1}{2}(N-n+1)}}{\Gamma\left(\frac{N-n+1}{2}\right)} \cdot \frac{{v_n}^{N-n-1}}{{v_{n-1}}^{N-n}} \cdot d\xi_{n1}, d\xi_{n2}, \dots\, d\xi_{nn} \quad \dots\dots\dots\dots(10).$$

(f) The quadratic co-ordinates of n points in space of N dimensions may be also regarded as co-ordinates of a single point in space of n sets, each of N dimensions. If the projections of the point upon the first $n-1$ spaces are fixed, the volume element corresponding to variations of the co-ordinates ξ_{n1}, ξ_{n2}, ... ξ_{nn} will be that found above.

(g) We now require the volume element corresponding to variations of all the co-ordinates of a point so specified in Nn dimensions. The component spaces of N dimensions each are defined by rectangular co-ordinates, hence the entire volume may be found by multiplying the volume element defined by the variations of ξ_{11}, ξ_{12}, ... $\xi_{n-1\,n-1}$ by that defined by the variations of ξ_{n1}, ξ_{n2}, ... ξ_{nn}. Consequently the volume element for variations of all the co-ordinates is found by multiplying together the volumes (equation (10)) got by putting n successively equal to 1, 2, ... n. The component volumes v_{n-r} appear in these expressions successively in the numerator and the denominator with equal indices, and therefore disappear from the product, thus we have finally for the volume element

$$\frac{\pi^{\frac{1}{4}n\,(2N-n+1)}}{\Gamma\left(\frac{N}{2}\right)\Gamma\left(\frac{N-1}{2}\right)\dots\Gamma\left(\frac{N-n+1}{2}\right)}\, {v_n}^{N-n-1}\, d\xi_{11}, d\xi_{12}, \dots\, d\xi_{nn} \quad \dots(11).$$

(h) Now write $N-1$ for N and let $\xi_{pq} = Na_{pq}$, to correspond with our earlier work. The above expression then becomes

$$\frac{\pi^{\frac{1}{4}n\,(2N-n-1)} \cdot N^{\frac{1}{2}n\,(N-1)}}{\Gamma\left(\frac{N-1}{2}\right)\Gamma\left(\frac{N-2}{2}\right)\dots\Gamma\left(\frac{N-n}{2}\right)} \begin{vmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \vdots & \vdots \quad \vdots \\ a_{n1} & a_{n2} \dots a_{nn} \end{vmatrix}^{\frac{N-n-2}{2}} da_{11} da_{12} \dots da_{nn},$$

and this multiplied by a density factor

$$(2\pi)^{-\frac{n}{2}(N-1)} \cdot \left(\frac{2}{N}\right)^{\frac{n}{2}(N-1)}$$

becomes

$$\frac{1}{(\sqrt{\pi})^{\frac{n(n-1)}{2}}\Gamma\left(\frac{N-1}{2}\right)\Gamma\left(\frac{N-2}{2}\right)\dots\Gamma\left(\frac{N-n}{2}\right)} \begin{vmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \vdots & \vdots \quad \vdots \\ a_{n1} & a_{n2} \dots a_{nn} \end{vmatrix}^{\frac{N-n-2}{2}} da_{11} da_{12} \dots da_{nn}$$

$$\dots\dots\dots\dots(12),$$

whence our general result (9) follows.

4. *Moment Coefficients of the Distribution.*

It is now possible to find the moments, up to any desired order, of the simultaneous distribution given by (9). A number of these moments will be identical with those found previously for the cases of one and two variates, and published by different writers. To illustrate for the case of three variates, let

$$F(abcfgh) = \frac{1}{\pi^{\frac{3}{2}}\,\Gamma\left(\frac{N-1}{2}\right)\Gamma\left(\frac{N-2}{2}\right)\Gamma\left(\frac{N-3}{2}\right)} \begin{vmatrix} A & H & G \\ H & B & F \\ G & F & C \end{vmatrix}^{\frac{N-1}{2}}$$

$$\times\, e^{-Aa-Bb-Cc-2Ff-2Gg-2Hh} \begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix}^{\frac{N-5}{2}} \qquad \text{.....(13).}$$

Then the generalised product moment coefficient of this distribution is given by

$$M_{jklmpq} = \iiint\!\!\iiint a^j b^k c^l f^m g^p h^q\, F(abcfgh)\, da\, db\, dc\, df\, dg\, dh,$$

where the limits of integration for a, b and c are 0 to ∞, and for f, g and h $-\sqrt{bc}$ to $+\sqrt{bc}$, $-\sqrt{ac}$ to $+\sqrt{ac}$ and $-\sqrt{ab}$ to $+\sqrt{ab}$ respectively. Now following the method of Romanovsky*, we define the function

$$\phi(\alpha\beta\gamma\lambda\mu\nu) = \int_0^\infty da \int_0^\infty db \int_0^\infty dc \iiint F(abcfgh)\, e^{a\alpha+b\beta+c\gamma+f\lambda+g\mu+h\nu}\, df\, dg\, dh$$

.....(14)

as the generating function of the product moments, which are given by

$$M_{jklmpq} = \left[\frac{\partial^{j+k+l+m+p+q}\,\phi(\alpha\beta\gamma\lambda\mu\nu)}{\partial\alpha^j \partial\beta^k \partial\gamma^l \partial\lambda^m \partial\mu^p \partial\nu^q}\right]_{\alpha=\beta=\gamma=\lambda=\mu=\nu=0} \qquad \text{.....(15).}$$

Since the total frequency volume in (8) is unity, we have

$$\int e^{-Aa-Bb-Cc-2Ff-2Gg-2Hh}\, dv = \text{const.} \begin{vmatrix} A & H & G \\ H & B & F \\ G & F & C \end{vmatrix}^{-\frac{N-1}{2}}$$

where the integration is taken throughout the volume. A parallel result is obtained for the integration in (14) simply by replacing A by $A-\alpha$, B by $B-\beta$, H by $H-\frac{1}{2}\nu$, etc., in the determinant $|ABC|$. What is then left to deal with is the fundamental frequency volume and we have

$$\phi(\alpha\beta\gamma\lambda\mu\nu) = \begin{vmatrix} A & H & G \\ H & B & F \\ G & F & C \end{vmatrix}^{\frac{N-1}{2}} \times \begin{vmatrix} A-\alpha & H-\frac{1}{2}\nu & G-\frac{1}{2}\mu \\ H-\frac{1}{2}\nu & B-\beta & F-\frac{1}{2}\lambda \\ G-\frac{1}{2}\mu & F-\frac{1}{2}\lambda & C-\gamma \end{vmatrix}^{-\frac{N-1}{2}} \qquad \text{...(16),}$$

a convenient expression from which to calculate the moments. It is of the same form when generalised for the case of n variates.

We shall use the notation $\mu\begin{pmatrix} j & q & p \\ & k & m \\ & & l \end{pmatrix}$,

corresponding to the determinant $\begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix}$,

* *Metron*, loc. cit. pp. 27, 32, 35—6.

to represent the general product moment coefficient of the jth power of s_1^2, kth power of s_2^2, lth power of s_3^2, mth power of $s_2 s_3 r_{23}$, pth power of $s_1 s_3 r_{13}$ and qth power of $s_1 s_2 r_{12}$. These moment coefficients are in all cases except the first calculated about the mean of the sample, and are expressed in terms of the following moment coefficients of the general population:

$$\mu_{200} = \sigma_1^2, \qquad \mu_{020} = \sigma_2^2, \qquad \mu_{002} = \sigma_3^2,$$

$$\mu_{110} = \sigma_1 \sigma_2 \rho_{12}, \quad \mu_{101} = \sigma_1 \sigma_3 \rho_{13}, \quad \mu_{011} = \sigma_2 \sigma_3 \rho_{23}.$$

In most cases there will be a number of similar results deduced from one another merely by a cyclical interchange of the suffixes, but one representative result for each class only will be stated. The following pages list every independent result up to the fourth order, and as far as eight variates. They therefore embody certain results which are already known. "Student"* first gave the one-variate moments up to the fourth order, while Soper†, to his degree of approximation, gave all the two-variate results, to the same order. His second and third order moments agree with the following, except for the terms in N, which he makes $\frac{1}{N}$ and $\frac{1}{N^2}$, in place of $\frac{N-1}{N^2}$ and $\frac{N-1}{N^3}$. The divergence of the fourth order results is wider, for the same reason. Isserlis‡ later gave the general 2nd order—4 variate, 3rd order—6 variate and 4th order—8 variate moments, sampling from a limited population. For an infinite normal population his results can be expressed in terms of the ρ's, and, of course, all the moments up to the fourth order can be deduced from his three results by associating the variates in all the possible ways. But his results were calculated about the general population value, and not, as in the following, about the mean of the sample.

5. *Derived Coefficients.*

We are now able to deduce certain other constants of the distribution curves specified by the moments cited. In particular we have for the Betas of a_{12}, i.e. of $s_1 s_2 r_{12}$,

$$a_{12}\beta_1 = \left\{\mu\begin{pmatrix}\cdot & 3\\ \cdot & \end{pmatrix}\right\}^2 \div \left\{\mu\begin{pmatrix}\cdot & 2\\ \cdot & \end{pmatrix}\right\}^3 = \frac{4}{N-1}\cdot\frac{\rho_{12}^2(3+\rho_{12}^2)^2}{(1+\rho_{12}^2)^3},$$

$$a_{12}\beta_2 = \left\{\mu\begin{pmatrix}\cdot & 4\\ \cdot & \end{pmatrix}\right\} \div \left\{\mu\begin{pmatrix}\cdot & 2\\ \cdot & \end{pmatrix}\right\}^2 = 3 + \frac{6}{N-1}\cdot\frac{1+6\rho_{12}^2+\rho_{12}^4}{(1+\rho_{12}^2)^2}.$$

These values show that when ρ_{12}, the correlation coefficient of the sampled population, is zero, the distribution of a_{12} is symmetrical, but not normal. For

$$\beta_2 = 3 + \frac{6}{N-1},$$

showing that as far as the first four moments the curve agrees with Type VII. On

* *Biometrika*, Vol. VI. 1908, p. 4.
† *Biometrika*, Vol. IX. 1913, pp. 103—104.
‡ *Biometrika*, Vol. XII. 1918, p. 183.

the other hand, when $\rho_{12}=1$ we have the Betas of s^2, the square of the standard deviation of the sample,

$$s^2\beta_1=\frac{8}{N-1}, \quad s^2\beta_2=3+\frac{12}{N-1},$$

the well-known Type III result*.

Since ${}_{a_{12}}\beta_2$ is always greater than 3, the curve of distribution is leptokurtic. We also have

$$2\beta_2-3\beta_1-6=\frac{12}{N-1}\cdot\frac{(1-\rho^2)^2}{(1+\rho^2)^3},$$

which is always positive, having a maximum value $\dfrac{12}{N-1}$. Also

$$k_2=\frac{\rho^2(3+\rho^2)^2}{4(1-\rho^2)^2}\left\{1+\frac{1}{N-1}\cdot\frac{1+6\rho^2+\rho^4}{(1+\rho^2)^2}\right\}^2\div\left\{1+\frac{1}{N-1}\cdot\frac{2+5\rho^2+8\rho^4+\rho^6}{(1+\rho^2)^3}\right\},$$

showing that the curve is of Type IV for small ρ, but beyond a value of about ·45 (according to the size of the sample) it becomes Type VI.

Certain correlation coefficients also follow at once from the moments given. It will be remembered that Pearson and Filon in *Phil. Trans.* Vol. 191 A, 1898, p. 242, gave $r_{\sigma_1\sigma_2}=\rho_{12}^2$ and $r_{\sigma_1 r_{13}}=\rho_{12}/\sqrt{2}$ approximately. They also gave approximate expressions for $r_{\sigma_1 . r_{23}}$, $r_{r_{12}.r_{13}}$ and $r_{r_{12}.r_{34}}$ (pp. 256, 259, 262). In our case the correlations we are able to deduce are those between the product moments themselves and are exact, thus

$$r_{a_{11}.a_{22}}=\left\{\mu\begin{pmatrix}1&\\&1\end{pmatrix}\right\}\div\left\{\mu\begin{pmatrix}2&\\&\end{pmatrix}\right\}^{\frac{1}{2}}\left\{\mu\begin{pmatrix}&\\&2\end{pmatrix}\right\}^{\frac{1}{2}}=\rho_{12}^2,$$

$$r_{a_{11}.a_{12}}=\left\{\mu\begin{pmatrix}1&1\\&\end{pmatrix}\right\}\div\left\{\mu\begin{pmatrix}2&\\&\end{pmatrix}\right\}^{\frac{1}{2}}\left\{\mu\begin{pmatrix}&\\&2\end{pmatrix}\right\}^{\frac{1}{2}}=\left\{\frac{2\rho_{12}^2}{1+\rho_{12}^2}\right\}^{\frac{1}{2}},$$

$$r_{a_{11}.a_{23}}=\left\{\mu\begin{pmatrix}1&&\\&1&\\&&\end{pmatrix}\right\}\div\left\{\mu\begin{pmatrix}2&&\\&&\\&&\end{pmatrix}\right\}^{\frac{1}{2}}\left\{\mu\begin{pmatrix}&&\\&&\\&&2\end{pmatrix}\right\}^{\frac{1}{2}}=\left\{\frac{2\rho_{12}^2\rho_{13}^2}{1+\rho_{23}^2}\right\}^{\frac{1}{2}},$$

$$r_{a_{23}.a_{13}}=\left\{\mu\begin{pmatrix}&1&\\&1&\\&&\end{pmatrix}\right\}\div\left\{\mu\begin{pmatrix}&2&\\&&\\&&\end{pmatrix}\right\}^{\frac{1}{2}}\left\{\mu\begin{pmatrix}&&\\&&\\&2&\end{pmatrix}\right\}^{\frac{1}{2}}=\frac{\rho_{12}+\rho_{13}\rho_{23}}{\{(1+\rho_{23}^2)(1+\rho_{13}^2)\}^{\frac{1}{2}}},$$

$$r_{a_{12}.a_{34}}=\left\{\mu\begin{pmatrix}1&&\\&&\\&&1\end{pmatrix}\right\}\div\left\{\mu\begin{pmatrix}2&&\\&&\\&&\end{pmatrix}\right\}^{\frac{1}{2}}\left\{\mu\begin{pmatrix}&&\\&&\\&&2\end{pmatrix}\right\}^{\frac{1}{2}}=\frac{\rho_{13}\rho_{24}+\rho_{14}\rho_{23}}{\{(1+\rho_{13}^2)(1+\rho_{24}^2)\}^{\frac{1}{2}}}.$$

My thanks are due to Dr R. A. Fisher, in whose laboratory this paper was written, and without whose critical help it would have been difficult to generalise the geometrical methods employed by him.

* "Student," *Biometrika*, Vol. VI. 1908, p. 4.

1ST ORDER:

1 *variate*

(1) $\mu(1) = \frac{N-1}{N}\sigma_1^2$.

2 *variates*

(2) $\mu\binom{\cdot}{\cdot}{}^{1} = \frac{N-1}{N}\sigma_1\sigma_2\rho_{12}$.

2ND ORDER:

1 *variate*

(3) $\mu(2) = 2\frac{N-1}{N^2}\sigma_1^4$.

2 *variates*

(4) $\mu\binom{1}{1} = 2\frac{N-1}{N^2}\sigma_1^2\sigma_2^2\rho_{12}^2$,

(5) $\mu\binom{\cdot}{\cdot}{}^{2} = \frac{N-1}{N^2}\sigma_1^2\sigma_2^2(1+\rho_{12}^2)$,

(6) $\mu\binom{1}{\cdot}{}^{1} = 2\frac{N-1}{N^2}\sigma_1^3\sigma_2\rho_{12}$.

3 *variates*

(7) $\mu\binom{1\ \cdot}{\cdot\ 1} = 2\frac{N-1}{N^2}\sigma_1^2\sigma_2\sigma_3\rho_{12}\rho_{13}$,

(8) $\mu\binom{\cdot\ \cdot}{\cdot\ 1}{}^{1} = \frac{N-1}{N^2}\sigma_1\sigma_2\sigma_3^2(\rho_{12}+\rho_{13}\rho_{23})$.

4 *variates*

(9) $\mu\binom{\cdot\ 1\ \cdot}{\cdot\ \cdot\ 1} = \frac{N-1}{N^2}\sigma_1\sigma_2\sigma_3\sigma_4(\rho_{13}\rho_{24}+\rho_{14}\rho_{23})$.

3RD ORDER:

1 *variate*

(10) $\mu(3) = 8\frac{N-1}{N^3}\sigma_1^6$.

2 *variates*

(11) $\mu\binom{2}{1} = 8\frac{N-1}{N^3}\sigma_1^4\sigma_2^2\rho_{12}^2$,

(12) $\mu\binom{\cdot}{\cdot}{}^{3} = 2\frac{N-1}{N^3}\sigma_1^3\sigma_2^3\rho_{12}(3+\rho_{12}^2)$,

(13) $\mu\binom{1}{\cdot}{}^{2} = 2\frac{N-1}{N^3}\sigma_1^4\sigma_2^2(1+3\rho_{12}^2)$,

(14) $\mu\binom{1}{1}{}^{1} = 4\frac{N-1}{N^3}\sigma_1^3\sigma_2^3\rho_{12}(1+\rho_{12}^2)$,

(15) $\mu\binom{2}{\cdot}{}^{1} = 8\frac{N-1}{N^3}\sigma_1^5\sigma_2\rho_{12}$.

3 *variates*

(16) $\mu\binom{2\ \cdot}{\cdot\ 1} = 8\frac{N-1}{N^3}\sigma_1^4\sigma_2\sigma_3\rho_{12}\rho_{13}$,

(17) $\mu\binom{1\ \cdot}{\cdot\ 2} = 2\frac{N-1}{N^3}\sigma_1^2\sigma_2^2\sigma_3^2(\rho_{12}^2+\rho_{13}^2+2\rho_{12}\rho_{13}\rho_{23})$,

(18) $\mu\binom{\cdot\ 1}{\cdot\ 2} = 2\frac{N-1}{N^3}\sigma_1\sigma_2^2\sigma_3^3(\rho_{13}+2\rho_{12}\rho_{23}+\rho_{13}\rho_{23}^2)$,

(19) $\mu\binom{1\ \cdot}{1\ \cdot} = 8\frac{N-1}{N^3}\sigma_1^2\sigma_2^2\sigma_3^2\rho_{12}\rho_{13}\rho_{23}$,

(20) $\mu\binom{1\,1\,1}{\cdot\ \cdot} = 2\frac{N-1}{N^3}\sigma_1^4\sigma_2\sigma_3(\rho_{23}+3\rho_{12}\rho_{13})$,

(21) $\mu\binom{\cdot\,1\,1}{\cdot\ 1} = \frac{N-1}{N^3}\sigma_1^2\sigma_2^2\sigma_3^2(1+\rho_{12}^2+\rho_{13}^2+\rho_{23}^2+4\rho_{12}\rho_{13}\rho_{23})$,

(22) $\mu\binom{1\ \cdot\ \cdot}{\cdot\ 1\ 1} = 4\frac{N-1}{N^3}\sigma_1^2\sigma_2^3\sigma_3\rho_{12}(\rho_{13}+\rho_{12}\rho_{23})$,

(23) $\mu\binom{1\ \cdot\ 1}{\cdot\ 1\ \cdot} = 2\frac{N-1}{N^3}\sigma_1^3\sigma_2\sigma_3^2(\rho_{12}+\rho_{13}\rho_{23}+2\rho_{12}\rho_{13}^2)$.

4 variates

(24) $\mu\begin{pmatrix}1 & \cdot & \cdot & \cdot \\ \cdot & \mathrm{i} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \mathrm{i}\end{pmatrix} = 4\frac{N-1}{N^3}\sigma_1^2\sigma_2^2\sigma_3\sigma_4\rho_{12}(\rho_{13}\rho_{24}+\rho_{14}\rho_{23}),$

(25) $\mu\begin{pmatrix}1 & \cdot & \cdot & \cdot \\ \cdot & \mathrm{i} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \mathrm{i}\end{pmatrix} = 2\frac{N-1}{N^3}\sigma_1^2\sigma_2\sigma_3^2\sigma_4[\rho_{12}\rho_{14}+\rho_{13}(\rho_{12}\rho_{34}+\rho_{13}\rho_{24}+\rho_{14}\rho_{23})],$

(26) $\mu\begin{pmatrix}\cdot & 1 & 1 & 1 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot\end{pmatrix} = 2\frac{N-1}{N^3}\sigma_1^3\sigma_2\sigma_3\sigma_4(\rho_{12}\rho_{34}+\rho_{13}\rho_{24}+\rho_{14}\rho_{23}+\rho_{12}\rho_{13}\rho_{14}),$

(27) $\mu\begin{pmatrix}1 & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \mathrm{i}\end{pmatrix} = 2\frac{N-1}{N^3}\sigma_1^3\sigma_2\sigma_3\sigma_4(\rho_{13}\rho_{24}+\rho_{14}\rho_{23}+2\rho_{12}\rho_{13}\rho_{14}),$

(28) $\mu\begin{pmatrix}\cdot & 2 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \mathrm{i}\end{pmatrix} = 2\frac{N-1}{N^3}\sigma_1^2\sigma_2^2\sigma_3\sigma_4[\rho_{13}\rho_{14}+\rho_{23}\rho_{24}+\rho_{12}(\rho_{13}\rho_{24}+\rho_{14}\rho_{23})],$

(29) $\mu\begin{pmatrix}\cdot & 1 & 1 & \cdot \\ \cdot & \cdot & \cdot & \mathrm{i} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot\end{pmatrix} = \frac{N-1}{N^3}\sigma_1^2\sigma_2^2\sigma_3\sigma_4[\rho_{34}+\rho_{13}\rho_{14}+\rho_{23}\rho_{24}+\rho_{12}(\rho_{13}\rho_{34}+\rho_{13}\rho_{24}+3\rho_{14}\rho_{23})].$

5 variates

(30) $\mu\begin{pmatrix}1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \mathrm{i} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \mathrm{i} & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot\end{pmatrix} = 2\frac{N-1}{N^3}\sigma_1^2\sigma_2\sigma_3\sigma_4\sigma_5[\rho_{12}(\rho_{14}\rho_{35}+\rho_{15}\rho_{34})+\rho_{13}(\rho_{14}\rho_{25}+\rho_{15}\rho_{24})],$

(31) $\mu\begin{pmatrix}\cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \mathrm{i} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \mathrm{i} & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot\end{pmatrix} = \frac{N-1}{N^3}\sigma_1\sigma_2^2\sigma_3\sigma_4\sigma_5[\rho_{14}\rho_{35}+\rho_{15}\rho_{34}+2\rho_{12}\rho_{24}\rho_{25}+\rho_{12}(\rho_{24}\rho_{35}+\rho_{25}\rho_{34})$
$+\rho_{23}(\rho_{14}\rho_{25}+\rho_{15}\rho_{24})].$

6 variates

(32) $\mu\begin{pmatrix}\cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \mathrm{i} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \mathrm{i} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\end{pmatrix} = \frac{N-1}{N^3}\sigma_1\sigma_2\sigma_3\sigma_4\sigma_5\sigma_6[\rho_{13}(\rho_{25}\rho_{46}+\rho_{26}\rho_{45})+\rho_{14}(\rho_{25}\rho_{36}+\rho_{26}\rho_{35})$
$+\rho_{15}(\rho_{23}\rho_{46}+\rho_{24}\rho_{36})+\rho_{16}(\rho_{23}\rho_{45}+\rho_{24}\rho_{35})].$

4TH ORDER:

1 variate

(33) $\mu(4) = 12\frac{(N-1)(N+3)}{N^4}\sigma_1^8.$

2 variates

(34) $\mu\begin{pmatrix}3 & \cdot \\ \cdot & \mathrm{i}\end{pmatrix} = 12\frac{(N-1)(N+3)}{N^4}\sigma_1^6\sigma_2^2\rho_{12}^2,$

(35) $\mu\begin{pmatrix}2 & \cdot \\ \cdot & 2\end{pmatrix} = 4\frac{N-1}{N^4}\sigma_1^4\sigma_2^4[N-1+8\rho_{12}^2+2(N+1)\rho_{12}^4],$

(36) $\mu\begin{pmatrix}2 & 1 \\ 1 & \cdot\end{pmatrix} = 4\frac{(N-1)(N+3)}{N^4}\sigma_1^5\sigma_2^3\rho_{12}(1+2\rho_{12}^2),$

(37) $\mu\begin{pmatrix}1 & 3 \\ \cdot & \cdot\end{pmatrix} = 6\frac{(N-1)(N+3)}{N^4}\sigma_1^5\sigma_2^3\rho_{12}(1+\rho_{12}^2),$

(38) $\mu\begin{pmatrix}3 & 1 \\ \cdot & \cdot\end{pmatrix} = 12\frac{(N-1)(N+3)}{N^4}\sigma_1^7\sigma_2\rho_{12},$

(39) $\mu\begin{pmatrix}2 & 2 \\ \cdot & \cdot\end{pmatrix} = 2\frac{(N-1)(N+3)}{N^4}\sigma_1^6\sigma_2^2(1+5\rho_{12}^2),$

(40) $\mu\begin{pmatrix}1 & 2 \\ 1 & \cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^4\sigma_2^4[2+(5N+11)\rho_{12}^2+(N+5)\rho_{12}^4],$

(41) $\mu\begin{pmatrix}\cdot & 4 \\ \cdot & \cdot\end{pmatrix} = 3\frac{N-1}{N^4}\sigma_1^4\sigma_2^4[(N+1)(1+\rho_{12}^4)+2(N+5)\rho_{12}^2].$

3 variates

(42) $\mu\begin{pmatrix}3 & \cdot & \cdot \\ \cdot & \mathrm{i} & \cdot \\ \cdot & \cdot & \cdot\end{pmatrix} = 12\,\frac{(N-1)(N+3)}{N^4}\,\sigma_1^6\sigma_2\sigma_3\rho_{12}\rho_{13},$

(43) $\mu\begin{pmatrix}2 & \cdot & \cdot \\ \mathrm{i} & \cdot & \cdot \\ 1 & \cdot & \cdot\end{pmatrix} = 4\,\frac{N-1}{N^4}\,\sigma_1^4\sigma_2^2\sigma_3^2[(N-1)\,\rho_{23}^2 + 8\rho_{12}\rho_{13}\rho_{23} + 2\,(N+1)\,\rho_{12}^2\rho_{13}^2],$

(44) $\mu\begin{pmatrix}2 & \cdot & \cdot \\ \mathrm{i} & \mathrm{i} & \cdot \\ \cdot & \cdot & \cdot\end{pmatrix} = 4\,\frac{N-1}{N^4}\,\sigma_1^4\sigma_2^3\sigma_3[(N-1)\,\rho_{23} + 4\rho_{12}\,(\rho_{13} + \rho_{12}\rho_{23}) + 2\,(N+1)\,\rho_{12}^3\rho_{13}],$

(45) $\mu\begin{pmatrix}2 & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & 2\end{pmatrix} = 2\,\frac{N-1}{N^4}\,\sigma_1^4\sigma_2^2\sigma_3^2[(N-1)\,(1+\rho_{23}^2) + 4\,(\rho_{12}^2 + \rho_{13}^2) + 8\rho_{12}\rho_{13}\rho_{23} + 4\,(N+1)\,\rho_{12}^2\rho_{13}^2],$

(46) $\mu\begin{pmatrix}2 & 1 & 1 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot\end{pmatrix} = 2\,\frac{(N-1)(N+3)}{N^4}\,\sigma_1^6\sigma_2\sigma_3\,(\rho_{23} + 5\rho_{12}\rho_{13}),$

(47) $\mu\begin{pmatrix}2 & \cdot & 1 \\ \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot\end{pmatrix} = 2\,\frac{(N-1)(N+3)}{N^4}\,\sigma_1^5\sigma_2\sigma_3^2\,(\rho_{12} + \rho_{13}\rho_{23} + 4\rho_{12}\rho_{13}^2),$

(48) $\mu\begin{pmatrix}1 & \cdot & \cdot \\ \mathrm{i} & \mathrm{i} & \cdot \\ 1 & \cdot & \cdot\end{pmatrix} = 4\,\frac{N-1}{N^4}\,\sigma_1^2\sigma_2^3\sigma_3^3[2\rho_{12}\rho_{13} + (N+1)\,\rho_{23}\,(\rho_{12}^2 + \rho_{13}^2) + (N+5)\,\rho_{12}\rho_{13}\rho_{23}^2],$

(49) $\mu\begin{pmatrix}1 & \cdot & \cdot \\ \cdot & 2 & \cdot \\ \cdot & \cdot & 2\end{pmatrix} = 2\,\frac{N-1}{N^4}\,\sigma_1^2\sigma_2^4\sigma_3^2[(N+3)\,\rho_{12}^2 + 2\rho_{13}^2 + 4\,(N+2)\,\rho_{12}\rho_{13}\rho_{23} + (N+5)\,\rho_{12}^2\rho_{23}^2],$

(50) $\mu\begin{pmatrix}1 & 1 & 1 \\ 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot\end{pmatrix} = 2\,\frac{N-1}{N^4}\,\sigma_1^4\sigma_2^3\sigma_3[2\rho_{23} + 2\,(N+2)\,\rho_{12}\rho_{13} + (3N+7)\,\rho_{12}^2\rho_{23} + (N+5)\,\rho_{12}^3\rho_{13}],$

(51) $\mu\begin{pmatrix}1 & \cdot & 1 \\ \mathrm{i} & \mathrm{i} & \cdot \\ \cdot & \cdot & \cdot\end{pmatrix} = 2\,\frac{N-1}{N^4}\,\sigma_1^3\sigma_2^3\sigma_3^2[2\,(\rho_{12} + N\rho_{13}\rho_{23}) + 4\rho_{12}\,(\rho_{13}^2 + \rho_{23}^2) + (N+1)\,\rho_{12}^3 + (3N+7)\,\rho_{12}^2\rho_{13}\rho_{23}],$

(52) $\mu\begin{pmatrix}1 & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & 3\end{pmatrix} = 6\,\frac{N-1}{N^4}\,\sigma_1^2\sigma_2^3\sigma_3^3[(N+1)\,\rho_{12}\rho_{13}\,(1+\rho_{23}^2) + 2\rho_{23}\,(\rho_{12}^2 + \rho_{13}^2)],$

(53) $\mu\begin{pmatrix}1 & 1 & 2 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot\end{pmatrix} = 2\,\frac{(N-1)(N+3)}{N^4}\,\sigma_1^5\sigma_2\sigma_3^2\,(\rho_{12} + 2\rho_{13}\rho_{23} + 3\rho_{12}\rho_{13}^2),$

(54) $\mu\begin{pmatrix}1 & \cdot & 1 \\ \cdot & 2 & \cdot \\ \cdot & \cdot & \cdot\end{pmatrix} = 2\,\frac{N-1}{N^4}\,\sigma_1^3\sigma_2^2\sigma_3^3[(N+1)\,\rho_{13}\,(1+\rho_{23}^2) + 4\rho_{12}\rho_{23} + 2\rho_{13}^3 + 2\,(N+2)\,\rho_{12}^2\rho_{13} + 2\,(N+3)\,\rho_{12}\rho_{13}^2\rho_{23}],$

(55) $\mu\begin{pmatrix}1 & 1 & 1 \\ \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot\end{pmatrix} = 2\,\frac{N-1}{N^4}\,\sigma_1^4\sigma_2^2\sigma_3^2[1 + \rho_{23}^2 + (N+2)\,(\rho_{12}^2 + \rho_{13}^2) + (3N+7)\,\rho_{12}\rho_{13}\rho_{23} + (N+5)\,\rho_{12}^2\rho_{13}^2],$

(56) $\mu\begin{pmatrix}\cdot & \cdot & 1 \\ \cdot & \cdot & 3 \\ \cdot & \cdot & \cdot\end{pmatrix} = 3\,\frac{N-1}{N^4}\,\sigma_1\sigma_2^3\sigma_3^4[(N+1)\,(\rho_{12} + \rho_{13}\rho_{23}^3) + (N+5)\,\rho_{23}\,(\rho_{13} + \rho_{12}\rho_{23})],$

(57) $\mu\begin{pmatrix}\cdot & \cdot & 2 \\ \cdot & \cdot & 2 \\ \cdot & \cdot & \cdot\end{pmatrix} = \frac{N-1}{N^4}\,\sigma_1^2\sigma_2^2\sigma_3^4[(N+1)\,(1+2\rho_{12}^2 + 3\rho_{13}^2\rho_{23}^2) + (N+5)\,(\rho_{13}^2 + \rho_{23}^2 + 4\rho_{12}\rho_{13}\rho_{23})],$

(58) $\mu\begin{pmatrix}\cdot & 1 & 1 \\ \cdot & \cdot & 2 \\ \cdot & \cdot & \cdot\end{pmatrix} = \frac{N-1}{N^4}\,\sigma_1^2\sigma_2^3\sigma_3^3[(N+5)\,\rho_{23} + (3N+7)\,\rho_{12}\rho_{13} + (N+1)\,\rho_{23}^3 + 2\,(N+3)\,\rho_{23}\,(\rho_{12}^2 + \rho_{13}^2) + (3N+11)\,\rho_{12}\rho_{13}\rho_{23}^2].$

4 variates

(59) $\mu\begin{pmatrix}2 & \cdot & \cdot & \cdot \\ \cdot & \mathrm{i} & \cdot & \mathrm{i} \\ \cdot & \cdot & \cdot & \cdot\end{pmatrix} = 4\,\frac{N-1}{N^4}\,\sigma_1^4\sigma_2^2\sigma_3\sigma_4[(N-1)\,\rho_{23}\rho_{24} + 4\rho_{12}\,(\rho_{13}\rho_{24} + \rho_{14}\rho_{23}) + 2\,(N+1)\,\rho_{12}^2\rho_{13}\rho_{14}],$

(60) $\mu\begin{pmatrix}2 & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \mathrm{i} \\ \cdot & \cdot & \cdot & \cdot\end{pmatrix} = 2\,\frac{(N-1)(N+3)}{N^4}\,\sigma_1^5\sigma_2\sigma_3\sigma_4\,(\rho_{13}\rho_{24} + \rho_{14}\rho_{23} + 4\rho_{12}\rho_{13}\rho_{14}),$

$$(61)\quad \mu\begin{pmatrix}2 & \cdot & \cdot\\ \cdot & 1 & 1\\ \cdot & \cdot & \cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^4\sigma_2^2\sigma_3\sigma_4[(N-1)(\rho_{34}+\rho_{23}\rho_{24})+4\rho_{12}(\rho_{12}\rho_{34}+\rho_{13}\rho_{21}+\rho_{14}\rho_{23}) + 4\rho_{13}\rho_{14}\{1+(N+1)\rho_{12}^2\}],$$

$$(62)\quad \mu\begin{pmatrix}1 & 1 & \cdot\\ 1 & \cdot & \cdot\\ 1 & \cdot & \cdot\end{pmatrix} = 4\frac{N-1}{N^4}\sigma_1^2\sigma_2^2\sigma_3^2\sigma_4^2[(N-1)(\rho_{12}^2\rho_{34}^2+\rho_{13}^2\rho_{24}^2+\rho_{14}^2\rho_{23}^2) + 4(\rho_{12}\rho_{13}\rho_{24}\rho_{34}+\rho_{12}\rho_{14}\rho_{23}\rho_{34}+\rho_{13}\rho_{14}\rho_{23}\rho_{24})],$$

$$(63)\quad \mu\begin{pmatrix}1 & \cdot & 1\\ 1 & \cdot & \cdot\\ \cdot & \cdot & \cdot\end{pmatrix} = 4\frac{N-1}{N^4}\sigma_1^3\sigma_2^2\sigma_3^2\sigma_4[(\rho_{12}\rho_{34}+\rho_{13}\rho_{24})\{2\rho_{23}+(N+1)\rho_{12}\rho_{13}\} + \rho_{14}\rho_{23}\{(N-1)\rho_{23}+4\rho_{12}\rho_{13}\}],$$

$$(64)\quad \mu\begin{pmatrix}1 & 1 & \cdot\\ 1 & \cdot & \cdot\\ \cdot & \cdot & 1\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^3\sigma_2^3\sigma_3\sigma_4[2(N+1)\rho_{12}(\rho_{13}\rho_{14}+\rho_{23}\rho_{24}) + (\rho_{13}\rho_{24}+\rho_{14}\rho_{23})\{2+(N+5)\rho_{12}^2\}],$$

$$(65)\quad \mu\begin{pmatrix}1 & \cdot & 1\\ 1 & 1 & \cdot\\ \cdot & \cdot & \cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^4\sigma_2^2\sigma_3\sigma_4[2\rho_{23}\rho_{24}+2(N+2)\rho_{12}(\rho_{13}\rho_{24}+\rho_{14}\rho_{23}) + \rho_{12}^2\{(N+3)\rho_{34}+(N+5)\rho_{13}\rho_{14}\}],$$

$$(66)\quad \mu\begin{pmatrix}1 & \cdot & 1\\ 1 & \cdot & \cdot\\ \cdot & 1 & \cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^3\sigma_2^3\sigma_3^2\sigma_4[2\rho_{23}(\rho_{12}\rho_{34}+\rho_{14}\rho_{23})+2\rho_{12}\rho_{24}(1+2\rho_{13}^2)+2N\rho_{13}\rho_{23}\rho_{24} + (N+1)\rho_{12}^2(\rho_{14}+\rho_{13}\rho_{34})+2(N+3)\rho_{12}\rho_{13}\rho_{14}\rho_{23}],$$

$$(67)\quad \mu\begin{pmatrix}1 & \cdot & \cdot\\ 1 & 1 & \cdot\\ \cdot & \cdot & 1\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^3\sigma_2^3\sigma_3\sigma_4[2\{\rho_{12}\rho_{34}+\rho_{13}\rho_{24}+(N-1)\rho_{14}\rho_{23}\}+4\rho_{12}\rho_{13}\rho_{14} + (N+1)\rho_{12}^2(\rho_{12}\rho_{34}+3\rho_{13}\rho_{24})+4\rho_{12}\rho_{23}(\rho_{24}+\rho_{12}\rho_{14})],$$

$$(68)\quad \mu\begin{pmatrix}1 & \cdot & \cdot\\ 1 & \cdot & \cdot\\ \cdot & 2 & \cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^2\sigma_2^2\sigma_3^2\sigma_4^2[(N-1)\rho_{12}^2(1+\rho_{34}^2)+2(\rho_{13}^2\rho_{24}^2+\rho_{14}^2\rho_{23}^2) + 4\rho_{12}\{\rho_{13}\rho_{23}+\rho_{14}\rho_{24}+\rho_{34}(\rho_{13}\rho_{24}+\rho_{14}\rho_{23})\}+4N\rho_{13}\rho_{14}\rho_{23}\rho_{24}],$$

$$(69)\quad \mu\begin{pmatrix}1 & 2 & \cdot\\ \cdot & \cdot & 1\\ \cdot & \cdot & \cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^4\sigma_2^2\sigma_3\sigma_4[(N+3)\rho_{13}\rho_{14}+2\rho_{23}\rho_{24}+2(N+2)\rho_{12}(\rho_{13}\rho_{24}+\rho_{14}\rho_{23}) + (N+5)\rho_{12}^2\rho_{13}\rho_{14}],$$

$$(70)\quad \mu\begin{pmatrix}1 & \cdot & 1\\ \cdot & 2 & \cdot\\ \cdot & \cdot & \cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^3\sigma_2^2\sigma_3^2\sigma_4[(N-1)\rho_{14}(1+\rho_{23}^2)+2(\rho_{12}\rho_{24}+\rho_{13}\rho_{34})+2\rho_{14}(\rho_{12}^2+\rho_{13}^2) + 2(\rho_{12}\rho_{34}+\rho_{13}\rho_{24})\{\rho_{23}+(N+1)\rho_{12}\rho_{13}\}+4\rho_{12}\rho_{13}\rho_{14}\rho_{23}],$$

$$(71)\quad \mu\begin{pmatrix}1 & \cdot & \cdot\\ \cdot & 2 & 1\\ \cdot & \cdot & \cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^2\sigma_2^3\sigma_3^2\sigma_4[2\rho_{24}(\rho_{12}^2+\rho_{13}^2)+4(\rho_{13}\rho_{14}\rho_{23}+\rho_{12}^2\rho_{23}\rho_{34}) + (N+1)\rho_{12}\{\rho_{14}(1+\rho_{23}^2)+2\rho_{13}(\rho_{34}+\rho_{23}\rho_{24})\}],$$

$$(72)\quad \mu\begin{pmatrix}1 & 1 & 1 & 1\\ \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot\end{pmatrix} = 2\frac{(N-1)(N+3)}{N^4}\sigma_1^5\sigma_2\sigma_3\sigma_4(\rho_{12}\rho_{34}+\rho_{13}\rho_{24}+\rho_{14}\rho_{23}+3\rho_{12}\rho_{13}\rho_{14}),$$

$$(73)\quad \mu\begin{pmatrix}1 & 1 & 1\\ \cdot & \cdot & 1\\ \cdot & \cdot & \cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^4\sigma_2^2\sigma_3\sigma_4[\rho_{34}+\rho_{23}\rho_{24}+(N+2)\{\rho_{13}(\rho_{14}+\rho_{12}\rho_{24})+\rho_{12}^2\rho_{34}\} + (N+5)\rho_{12}^2\rho_{13}\rho_{14}+(2N+5)\rho_{12}\rho_{14}\rho_{23}],$$

$$(74)\quad \mu\begin{pmatrix}1 & 1 & \cdot\\ \cdot & 1 & 1\\ \cdot & \cdot & \cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^3\sigma_2^3\sigma_3\sigma_4[(N+1)\rho_{12}(\rho_{34}+\rho_{23}\rho_{24})+2(N+2)\rho_{12}\rho_{13}\rho_{14} + 2(\rho_{13}\rho_{24}+\rho_{14}\rho_{23}+\rho_{12}^2\rho_{34})+(N+3)\rho_{12}^2(\rho_{13}\rho_{24}+\rho_{14}\rho_{23})],$$

$$(75)\quad \mu\begin{pmatrix}1 & 1 & \cdot\\ \cdot & 1 & \cdot\\ \cdot & \cdot & 1\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^3\sigma_2^2\sigma_3^2\sigma_4[\rho_{13}(\rho_{34}+3\rho_{23}\rho_{24})+\rho_{14}\{1+2\rho_{12}^2+\rho_{23}^2+(N+1)\rho_{13}^2 + 2(N+2)\rho_{12}\rho_{13}\rho_{23}\}+N\rho_{12}(\rho_{24}+\rho_{23}\rho_{34})+\rho_{12}\rho_{13}\{2\rho_{12}\rho_{34}+(N+3)\rho_{13}\rho_{24}\}],$$

$$(76)\quad \mu\begin{pmatrix}1 & \cdot & \mathrm{i} & \mathrm{i} \\ \cdot & \cdot & \cdot & \mathrm{i} \\ \cdot & \cdot & \cdot & \cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^3\sigma_2^3\sigma_3^3\sigma_4^3[\rho_{12}^2+\rho_{13}^2+\rho_{14}^2+\rho_{12}^2\rho_{34}^2+\rho_{13}^2\rho_{24}^2+\rho_{14}^2\rho_{23}^2$$
$$+\rho_{12}\rho_{13}\{(N+1)\rho_{23}+(N+3)\rho_{24}\rho_{34}\}+\rho_{12}\rho_{14}\{(N+1)\rho_{24}+(N+3)\rho_{23}\rho_{34}\}$$
$$+\rho_{13}\rho_{14}\{(N+1)\rho_{34}+(N+3)\rho_{23}\rho_{24}\}],$$

$$(77)\quad \mu\begin{pmatrix}\cdot & 3 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \mathrm{i} \\ \cdot & \cdot & \cdot & \cdot\end{pmatrix} = 3\frac{N-1}{N^4}\sigma_1^3\sigma_2^3\sigma_3^3\sigma_4[4\rho_{12}(\rho_{13}\rho_{14}+\rho_{23}\rho_{24})+(N+1)(1+\rho_{12}^2)(\rho_{13}\rho_{24}+\rho_{14}\rho_{23})],$$

$$(78)\quad \mu\begin{pmatrix}\cdot & 2 & 1 & 1 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot\end{pmatrix} = \frac{N-1}{N^4}\sigma_1^4\sigma_2^2\sigma_3^2\sigma_4[(N+1)(\rho_{34}+2\rho_{23}\rho_{24}+3\rho_{12}^2\rho_{13}\rho_{14})$$
$$+(N+5)(\rho_{13}\rho_{14}+\rho_{12}^2\rho_{34}+2\rho_{12}\rho_{13}\rho_{24}+2\rho_{12}\rho_{23}\rho_{14})],$$

$$(79)\quad \mu\begin{pmatrix}\cdot & 2 & 1 & \cdot \\ \cdot & \cdot & \cdot & \mathrm{i} \\ \cdot & \cdot & \cdot & \cdot\end{pmatrix} = \frac{N-1}{N^4}\sigma_1^3\sigma_2^3\sigma_3^3\sigma_4[(N+5)\rho_{12}\rho_{34}+4\rho_{13}\rho_{24}$$
$$+(N+1)(3\rho_{14}\rho_{23}+2\rho_{12}^2\rho_{13}\rho_{24}+\rho_{12}^3\rho_{34})$$
$$+2(N+3)\rho_{12}(\rho_{13}\rho_{14}+\rho_{23}\rho_{24})+(N+9)\rho_{12}^2\rho_{14}\rho_{23}],$$

$$(80)\quad \mu\begin{pmatrix}\cdot & 2 & 1 & \cdot \\ \cdot & \cdot & \cdot & \mathrm{i} \\ \cdot & \cdot & \cdot & \cdot\end{pmatrix} = \frac{N-1}{N^4}\sigma_1^3\sigma_2^2\sigma_3^2\sigma_4[4\rho_{12}(\rho_{24}+\rho_{23}\rho_{34})+4\rho_{13}^2\rho_{14}$$
$$+(N+1)\{(1+\rho_{12}^2)(\rho_{14}+\rho_{13}\rho_{34})+2(\rho_{14}\rho_{23}^2+\rho_{12}\rho_{13}^2\rho_{24})\}$$
$$+2(N+3)\rho_{13}\rho_{23}\rho_{24}+2(N+5)\rho_{12}\rho_{13}\rho_{14}\rho_{23}],$$

$$(81)\quad \mu\begin{pmatrix}\cdot & 2 & \cdot & \cdot \\ \cdot & \cdot & \cdot & 2 \\ \cdot & \cdot & \cdot & \cdot\end{pmatrix} = \frac{N-1}{N^4}\sigma_1^2\sigma_2^2\sigma_3^2\sigma_4^2[(N-1)(1+\rho_{12}^2)(1+\rho_{34}^2)+2(\rho_{13}^2+\rho_{14}^2+\rho_{23}^2+\rho_{24}^2)$$
$$+2N(\rho_{13}^2\rho_{24}^2+\rho_{14}^2\rho_{23}^2)+4\{\rho_{12}\rho_{13}\rho_{23}+\rho_{12}\rho_{14}\rho_{24}+\rho_{13}\rho_{14}\rho_{34}$$
$$+\rho_{23}\rho_{24}\rho_{34}+\rho_{12}\rho_{34}(\rho_{13}\rho_{24}+\rho_{14}\rho_{23})+(N+2)\rho_{13}\rho_{14}\rho_{23}\rho_{24}\}],$$

$$(82)\quad \mu\begin{pmatrix}\cdot & 1 & 1 & 1 \\ \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot\end{pmatrix} = \frac{N-1}{N^4}\sigma_1^3\sigma_2^2\sigma_3^2\sigma_4[\rho_{14}\{2(1+\rho_{23}^2)+(N+1)(\rho_{12}^2+\rho_{13}^2)\}$$
$$+2(N+2)\rho_{23}(\rho_{12}\rho_{34}+\rho_{13}\rho_{24})+(N+3)(\rho_{12}\rho_{24}+\rho_{13}\rho_{34}+2\rho_{12}\rho_{13}\rho_{14}\rho_{23})$$
$$+(N+5)\rho_{12}\rho_{13}(\rho_{12}\rho_{34}+\rho_{13}\rho_{24})],$$

$$(83)\quad \mu\begin{pmatrix}\cdot & 1 & 1 & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & 1 & \cdot & \cdot\end{pmatrix} = \frac{N-1}{N^4}\sigma_1^2\sigma_2^2\sigma_3^2\sigma_4^2[1+\rho_{12}^2+\rho_{13}^2+\rho_{24}^2+\rho_{34}^2+N(\rho_{14}^2+\rho_{23}^2)$$
$$+\rho_{12}^2\rho_{34}^2+\rho_{13}^2\rho_{24}^2+(N+2)\rho_{14}^2\rho_{23}^2$$
$$+(N+3)(\rho_{12}\rho_{13}\rho_{23}+\rho_{12}\rho_{14}\rho_{24}+\rho_{13}\rho_{14}\rho_{34}+\rho_{23}\rho_{24}\rho_{34})$$
$$+(N+7)\rho_{14}\rho_{23}(\rho_{12}\rho_{24}+\rho_{13}\rho_{24})+(3N+1)\rho_{12}\rho_{13}\rho_{24}\rho_{34}].$$

5 variates

$$(84)\quad \mu\begin{pmatrix}2 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \mathrm{i} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \mathrm{i} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^4\sigma_2\sigma_3\sigma_4\sigma_5[(N-1)(\rho_{24}\rho_{35}+\rho_{25}\rho_{34})+4\rho_{12}(\rho_{14}\rho_{35}+\rho_{15}\rho_{34})$$
$$+4\rho_{13}(\rho_{14}\rho_{25}+\rho_{15}\rho_{24})+4(N+1)\rho_{12}\rho_{13}\rho_{14}\rho_{15}],$$

$$(85)\quad \mu\begin{pmatrix}1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \mathrm{i} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \mathrm{i} & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot\end{pmatrix} = 4\frac{N-1}{N^4}\sigma_1^2\sigma_2^2\sigma_3^2\sigma_4\sigma_5[(N-1)(\rho_{12}^2\rho_{34}\rho_{35}+\rho_{13}^2\rho_{24}\rho_{25}+\rho_{23}^2\rho_{14}\rho_{15})$$
$$+2\rho_{12}\rho_{13}(\rho_{24}\rho_{35}+\rho_{25}\rho_{34})+2\rho_{12}\rho_{23}(\rho_{14}\rho_{35}+\rho_{15}\rho_{34})$$
$$+2\rho_{13}\rho_{23}(\rho_{14}\rho_{25}+\rho_{15}\rho_{24})],$$

$$(86)\quad \mu\begin{pmatrix}1 & \mathrm{i} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \mathrm{i} & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^2\sigma_2^3\sigma_3^3\sigma_4\sigma_5[2\rho_{12}(\rho_{14}\rho_{35}+\rho_{15}\rho_{34})+2(\rho_{13}+2\rho_{12}\rho_{23})(\rho_{14}\rho_{25}+\rho_{15}\rho_{24})$$
$$+2(N-1)\rho_{14}\rho_{15}\rho_{23}+(N+1)\rho_{12}\{\rho_{12}(\rho_{24}\rho_{35}+\rho_{25}\rho_{34})+2\rho_{13}\rho_{24}\rho_{25}\}],$$

$$(87)\quad \mu\begin{pmatrix}1&\cdot&\cdot&\cdot\\ \cdot&\mathrm{i}&\mathrm{i}&\cdot\\ \cdot&\cdot&\cdot&\cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^2\sigma_2^2\sigma_3^2\sigma_4\sigma_5[(N-1)\rho_{12}^2(\rho_{45}+\rho_{34}\rho_{35}) + 2(\rho_{13}^2\rho_{24}\rho_{25}+\rho_{23}^2\rho_{14}\rho_{15}) + 2\rho_{12}\rho_{13}(\rho_{24}\rho_{35}+\rho_{25}\rho_{34}) + 2\rho_{12}\rho_{23}(2\rho_{13}\rho_{45}+\rho_{14}\rho_{35}+\rho_{15}\rho_{34}) + 2(\rho_{12}+N\rho_{13}\rho_{23})(\rho_{14}\rho_{25}+\rho_{15}\rho_{24})],$$

$$(88)\quad \mu\begin{pmatrix}1&\cdot&\cdot&\cdot\\ \cdot&2&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\mathrm{i}&\cdot\\ \cdot&\cdot&\cdot&\cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^2\sigma_2^2\sigma_3^2\sigma_4\sigma_5[(N-1)\rho_{14}\rho_{15}(1+\rho_{22}^2) + 2(\rho_{12}^2\rho_{34}\rho_{35}+\rho_{13}^2\rho_{24}\rho_{25}) + 2N\rho_{12}\rho_{13}(\rho_{24}\rho_{35}+\rho_{25}\rho_{34}) + 2\rho_{12}\rho_{14}(\rho_{25}+\rho_{23}\rho_{35}) + 2\rho_{12}\rho_{15}(\rho_{24}+\rho_{23}\rho_{34}) + 2\rho_{13}\rho_{14}(\rho_{35}+\rho_{23}\rho_{25}) + 2\rho_{13}\rho_{15}(\rho_{34}+\rho_{23}\rho_{24})],$$

$$(89)\quad \mu\begin{pmatrix}1&1&1&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\mathrm{i}&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^4\sigma_2\sigma_3\sigma_4\sigma_5[\rho_{24}\rho_{35}+\rho_{25}\rho_{34}+2\rho_{12}\rho_{13}\rho_{14}\rho_{15} + (N+2)\{\rho_{12}(\rho_{14}\rho_{35}+\rho_{15}\rho_{34})+\rho_{13}(\rho_{14}\rho_{25}+\rho_{15}\rho_{24})\} + (N+3)\rho_{14}\rho_{15}(\rho_{23}+\rho_{12}\rho_{13})],$$

$$(90)\quad \mu\begin{pmatrix}1&1&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\mathrm{i}&\mathrm{i}&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^3\sigma_2\sigma_3^2\sigma_4\sigma_5[\rho_{14}\rho_{25}+\rho_{15}\rho_{24}+2\rho_{12}\rho_{14}\rho_{15}+(N-1)\rho_{12}(\rho_{45}+\rho_{34}\rho_{35}) + \rho_{13}(\rho_{23}\rho_{45}+\rho_{24}\rho_{35}+\rho_{25}\rho_{34})+(\rho_{23}+2\rho_{12}\rho_{13})(\rho_{13}\rho_{45}+\rho_{14}\rho_{35}+\rho_{15}\rho_{34}) + (N+1)\{\rho_{13}^2(\rho_{14}\rho_{25}+\rho_{15}\rho_{24})+2\rho_{13}\rho_{14}\rho_{15}\rho_{23}\}],$$

$$(91)\quad \mu\begin{pmatrix}1&\cdot&1&\cdot&\cdot\\ \cdot&\cdot&\mathrm{i}&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\mathrm{i}&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^3\sigma_2\sigma_3^2\sigma_4\sigma_5[(1+2\rho_{13}^2)(\rho_{14}\rho_{25}+\rho_{15}\rho_{24}) + 2\rho_{12}\rho_{34}\rho_{35}+\rho_{23}(\rho_{14}\rho_{35}+\rho_{15}\rho_{34})+N\rho_{13}(\rho_{24}\rho_{35}+\rho_{25}\rho_{34}) + (N+3)\rho_{12}\rho_{13}(\rho_{14}\rho_{35}+\rho_{15}\rho_{34})+(N+1)\rho_{14}\rho_{15}(\rho_{12}+\rho_{13}\rho_{23})],$$

$$(92)\quad \mu\begin{pmatrix}1&\cdot&\cdot&\cdot\\ \cdot&\mathrm{i}&\cdot&\cdot\\ \cdot&\mathrm{i}&\mathrm{i}&\cdot\\ \cdot&\cdot&\cdot&\cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^3\sigma_2\sigma_3^2\sigma_4\sigma_5[(N-1)\rho_{12}(\rho_{45}+\rho_{34}\rho_{35})+(N+1)\{\rho_{13}^2(\rho_{14}\rho_{25}+\rho_{15}\rho_{24}) + 2\rho_{13}\rho_{14}\rho_{15}\rho_{23}\}+\rho_{14}\rho_{25}+\rho_{15}\rho_{24}+2\rho_{12}\rho_{14}\rho_{15}+2\rho_{12}\rho_{13}(\rho_{14}\rho_{35}+\rho_{15}\rho_{34}) + 2\rho_{13}\rho_{45}(\rho_{23}+\rho_{12}\rho_{13})+\rho_{34}(\rho_{13}\rho_{25}+\rho_{15}\rho_{23})+\rho_{35}(\rho_{13}\rho_{24}+\rho_{14}\rho_{23})],$$

$$(93)\quad \mu\begin{pmatrix}1&\cdot&\cdot&\cdot\\ \cdot&\mathrm{i}&\cdot&\cdot\\ \cdot&\cdot&\mathrm{i}&\cdot\\ \cdot&\cdot&1&\cdot\\ \cdot&\cdot&\cdot&\cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^2\sigma_2\sigma_3^2\sigma_4\sigma_5^2[\rho_{12}\rho_{14}(1+\rho_{35}^2)+\rho_{12}\rho_{15}\rho_{45}+\rho_{24}(\rho_{13}^2+\rho_{15}^2) + \rho_{13}\rho_{23}(\rho_{14}+\rho_{15}\rho_{45})+\rho_{13}\rho_{25}(\rho_{13}\rho_{45}+3\rho_{14}\rho_{35})+\rho_{15}\rho_{34}(\rho_{15}\rho_{23}+3\rho_{12}\rho_{35}) + N\{\rho_{12}\rho_{13}(\rho_{34}+\rho_{35}\rho_{45})+\rho_{14}\rho_{15}(\rho_{25}+\rho_{23}\rho_{35})\}+(N+1)\rho_{13}\rho_{15}\rho_{24}\rho_{35} + (N+2)\rho_{13}\rho_{15}\rho_{25}\rho_{34}],$$

$$(94)\quad \mu\begin{pmatrix}\cdot&2&1&\cdot\\ \cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\mathrm{i}&\cdot\\ \cdot&\cdot&\cdot&\cdot\end{pmatrix} = \frac{N-1}{N^4}\sigma_1^2\sigma_2^2\sigma_3\sigma_4\sigma_5[4\rho_{13}(\rho_{24}\rho_{35}+\rho_{25}\rho_{34})+4\rho_{13}\rho_{24}\rho_{25}+4\rho_{14}\rho_{15}(\rho_{13}+2\rho_{12}\rho_{23}) + (N+1)\{(1+\rho_{12}^2)(\rho_{14}\rho_{35}+\rho_{15}\rho_{34})+2(\rho_{23}+\rho_{12}\rho_{13})(\rho_{14}\rho_{25}+\rho_{15}\rho_{24})\}],$$

$$(95)\quad \mu\begin{pmatrix}\cdot&2&\cdot&\cdot\\ \cdot&\cdot&\mathrm{i}&\mathrm{i}\\ \cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot\end{pmatrix} = \frac{N-1}{N^4}\sigma_1^2\sigma_2^2\sigma_3^2\sigma_4\sigma_5[(N-1)(1+\rho_{12}^2)(\rho_{45}+\rho_{34}\rho_{35})+2N(\rho_{13}^2\rho_{24}\rho_{25}+\rho_{23}^2\rho_{14}\rho_{15}) + 2(N+2)\rho_{13}\rho_{23}(\rho_{14}\rho_{25}+\rho_{15}\rho_{24})+2(\rho_{14}\rho_{15}+\rho_{24}\rho_{25})+2\rho_{45}(\rho_{13}^2+\rho_{23}^2) + 2\rho_{12}(\rho_{14}\rho_{25}+\rho_{15}\rho_{24})+2\rho_{13}(\rho_{14}\rho_{35}+\rho_{15}\rho_{34})+2\rho_{23}(\rho_{24}\rho_{35}+\rho_{25}\rho_{34}) + 2\rho_{12}\rho_{13}(\rho_{23}\rho_{45}+\rho_{24}\rho_{35}+\rho_{25}\rho_{34})+2\rho_{12}\rho_{23}(\rho_{13}\rho_{45}+\rho_{14}\rho_{25}+\rho_{15}\rho_{34})],$$

$$(96)\quad \mu\begin{pmatrix}\cdot&1&1&1&1\\ \cdot&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot\end{pmatrix} = \frac{N-1}{N^4}\sigma_1^4\sigma_2\sigma_3\sigma_4\sigma_5[(N+1)(\rho_{23}\rho_{45}+\rho_{24}\rho_{35}+\rho_{25}\rho_{34}+3\rho_{12}\rho_{13}\rho_{14}\rho_{15}) + (N+5)\{\rho_{12}(\rho_{13}\rho_{45}+\rho_{14}\rho_{35})+\rho_{14}(\rho_{13}\rho_{25}+\rho_{15}\rho_{23})+\rho_{15}(\rho_{12}\rho_{34}+\rho_{13}\rho_{24})\}],$$

$$(97)\quad \mu\begin{pmatrix}\cdot 111\cdot\\ \cdot\cdot\cdot 1\\ \cdot\cdot 1\cdot\\ \cdot\cdot\cdot\cdot\end{pmatrix} = \frac{N-1}{N^4}\sigma_1^3\sigma_2\sigma_3^2\sigma_4\sigma_5[2\rho_{12}(\rho_{45}+\rho_{34}\rho_{35})+2\rho_{14}(\rho_{25}+\rho_{23}\rho_{35})+4\rho_{13}^2\rho_{15}\rho_{24}$$
$$+(N+1)\{\rho_{13}^2(\rho_{12}\rho_{45}+\rho_{14}\rho_{25})+2\rho_{15}\rho_{23}\rho_{34}+(\rho_{15}+\rho_{13}\rho_{35})(\rho_{24}+\rho_{12}\rho_{14})\}$$
$$+(N+3)\rho_{13}(\rho_{23}\rho_{45}+\rho_{25}\rho_{34})+(N+5)\rho_{13}\rho_{15}(\rho_{13}\rho_{34}+\rho_{14}\rho_{23})],$$

$$(98)\quad \mu\begin{pmatrix}\cdot 11\cdot\cdot\\ \cdot 1\cdot\cdot\\ \cdot\cdot\cdot 1\\ \cdot\cdot\cdot\cdot\end{pmatrix} = \frac{N-1}{N^4}\sigma_1^2\sigma_2^2\sigma_3^2\sigma_4\sigma_5[2\{\rho_{14}\rho_{15}(1+\rho_{23}^2)+\rho_{24}\rho_{25}(1+\rho_{13}^2)+\rho_{34}\rho_{35}(1+\rho_{12}^2)\}$$
$$+(\rho_{14}\rho_{25}+\rho_{15}\rho_{24})\{(N+1)\rho_{12}+(N+3)\rho_{13}\rho_{23}\}$$
$$+(\rho_{14}\rho_{35}+\rho_{15}\rho_{34})\{(N+1)\rho_{13}+(N+3)\rho_{12}\rho_{23}\}$$
$$+(\rho_{24}\rho_{35}+\rho_{25}\rho_{34})\{(N+1)\rho_{23}+(N+3)\rho_{12}\rho_{13}\}],$$

$$(99)\quad \mu\begin{pmatrix}\cdot 11\cdot\\ \cdot\cdot 1\cdot\\ \cdot\cdot\cdot 1\\ \cdot\cdot\cdot\cdot\end{pmatrix} = \frac{N-1}{N^4}\sigma_1^2\sigma_2^2\sigma_3^2\sigma_4\sigma_5[\rho_{45}(1+\rho_{12}^2+\rho_{13}^2+N\rho_{23}^2)+\rho_{14}\rho_{15}\{N+(N+2)\rho_{23}^2\}$$
$$+\rho_{24}\rho_{25}(1+\rho_{13}^2)+\rho_{34}\rho_{35}(1+\rho_{12}^2)+\rho_{23}\rho_{24}\rho_{35}+2\rho_{13}\rho_{14}\rho_{23}\rho_{25}$$
$$+3\rho_{12}\rho_{14}(\rho_{25}+\rho_{23}\rho_{35})+3\rho_{13}\rho_{15}(\rho_{34}+\rho_{23}\rho_{24})+3\rho_{12}\rho_{23}(\rho_{13}\rho_{45}+\rho_{15}\rho_{34})$$
$$+N\{\rho_{12}\rho_{13}(\rho_{23}\rho_{45}+\rho_{24}\rho_{35}+\rho_{25}\rho_{34})+\rho_{12}\rho_{15}\rho_{24}+\rho_{13}\rho_{14}\rho_{35}\}$$
$$+(N+1)\rho_{12}\rho_{34}(\rho_{13}\rho_{25}+\rho_{15}\rho_{23})+(N+2)\rho_{23}\rho_{25}(\rho_{34}+\rho_{13}\rho_{14})],$$

$$(100)\quad \mu\begin{pmatrix}\cdot 11\cdot\\ \cdot\cdot 1\cdot\\ \cdot\cdot\cdot 1\\ \cdot\cdot\cdot\cdot\end{pmatrix} = \frac{N-1}{N^4}\sigma_1^2\sigma_2^3\sigma_3\sigma_4^3\sigma_5[(1+\rho_{12}^2)(\rho_{35}+\rho_{34}\rho_{45})+(\rho_{13}+3\rho_{12}\rho_{23})(\rho_{15}+\rho_{14}\rho_{45})$$
$$+\rho_{24}^2(\rho_{35}+\rho_{13}\rho_{15})+3\rho_{24}(\rho_{25}\rho_{34}+\rho_{13}\rho_{14}\rho_{25})$$
$$+N\{\rho_{14}^2\rho_{35}+(\rho_{23}+\rho_{12}\rho_{13})(\rho_{25}+\rho_{24}\rho_{45})\}+(2N+1)\rho_{12}\rho_{15}\rho_{24}\rho_{34}$$
$$+(N+2)\rho_{14}(\rho_{15}\rho_{34}+\rho_{14}\rho_{23}\rho_{25})+(N+3)\rho_{12}\rho_{14}\rho_{24}\rho_{35}$$
$$+(N+4)\rho_{14}(\rho_{12}\rho_{25}\rho_{34}+\rho_{15}\rho_{23}\rho_{24})].$$

6 variates

$$(101)\quad \mu\begin{pmatrix}1\cdot\cdot\cdot\cdot\cdot\\ 1\cdot\cdot\cdot\cdot\cdot\\ \cdot 1\cdot\cdot\cdot\cdot\\ \cdot\cdot\cdot\cdot 1\\ \cdot\cdot\cdot\cdot\cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^2\sigma_2^2\sigma_3\sigma_4\sigma_5\sigma_6[(N-1)\{\rho_{12}^2(\rho_{35}\rho_{46}+\rho_{36}\rho_{45})$$
$$+2\rho_{13}\rho_{14}\rho_{25}\rho_{26}+2\rho_{15}\rho_{16}\rho_{23}\rho_{24}\}+2(\rho_{13}\rho_{24}+\rho_{14}\rho_{23})(\rho_{15}\rho_{26}+\rho_{16}\rho_{25})$$
$$+2\rho_{12}\rho_{13}(\rho_{25}\rho_{46}+\rho_{26}\rho_{45})+2\rho_{12}\rho_{14}(\rho_{25}\rho_{36}+\rho_{26}\rho_{35})$$
$$+2\rho_{12}\rho_{15}(\rho_{23}\rho_{46}+\rho_{24}\rho_{36})+2\rho_{12}\rho_{16}(\rho_{23}\rho_{45}+\rho_{24}\rho_{35})],$$

$$(102)\quad \mu\begin{pmatrix}11\cdot\cdot\cdot\cdot\\ \cdot\cdot\cdot\cdot\cdot\\ \cdot 1\cdot\cdot\cdot\cdot\\ \cdot\cdot\cdot 1\cdot\\ \cdot\cdot\cdot\cdot 1\\ \cdot\cdot\cdot\cdot\cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^3\sigma_2\sigma_3\sigma_4\sigma_5\sigma_6[(N-1)\rho_{12}(\rho_{35}\rho_{46}+\rho_{36}\rho_{45})+\rho_{13}(\rho_{25}\rho_{46}+\rho_{26}\rho_{45})$$
$$+\rho_{14}(\rho_{25}\rho_{36}+\rho_{26}\rho_{35})+\rho_{15}(\rho_{23}\rho_{46}+\rho_{24}\rho_{36})+\rho_{16}(\rho_{23}\rho_{45}+\rho_{24}\rho_{35})$$
$$+2\rho_{12}\rho_{13}(\rho_{15}\rho_{46}+\rho_{16}\rho_{45})+2\rho_{12}\rho_{14}(\rho_{15}\rho_{36}+\rho_{16}\rho_{35})$$
$$+(N+1)\{\rho_{13}\rho_{15}(\rho_{14}\rho_{26}+\rho_{16}\rho_{24})+\rho_{14}\rho_{16}(\rho_{13}\rho_{25}+\rho_{15}\rho_{23})\}],$$

$$(103)\quad \mu\begin{pmatrix}1\cdot\cdot\cdot\cdot\cdot\\ \cdot 1\cdot\cdot\cdot\cdot\\ \cdot\cdot 1\cdot\cdot\cdot\\ \cdot\cdot\cdot\cdot 1\\ \cdot\cdot\cdot\cdot\cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^2\sigma_2\sigma_3^2\sigma_4\sigma_5\sigma_6[\rho_{12}(\rho_{15}\rho_{46}+\rho_{16}\rho_{45})+\rho_{14}(\rho_{15}\rho_{26}+\rho_{16}\rho_{25})$$
$$+2\rho_{12}\rho_{14}\rho_{35}\rho_{36}+2\rho_{13}\rho_{15}\rho_{24}\rho_{36}+2\rho_{13}\rho_{16}\rho_{24}\rho_{35}+\rho_{13}^2(\rho_{25}\rho_{46}+\rho_{26}\rho_{45})$$
$$+N\{\rho_{12}\rho_{13}(\rho_{35}\rho_{46}+\rho_{36}\rho_{45})+\rho_{13}\rho_{14}(\rho_{25}\rho_{36}+\rho_{26}\rho_{35})\}+\rho_{13}\rho_{15}(\rho_{23}\rho_{46}+\rho_{26}\rho_{34})$$
$$+\rho_{13}\rho_{16}(\rho_{23}\rho_{45}+\rho_{25}\rho_{34})+(N-1)\rho_{15}\rho_{16}(\rho_{34}+\rho_{23}\rho_{34})$$
$$+(\rho_{12}\rho_{34}+\rho_{14}\rho_{23})(\rho_{15}\rho_{36}+\rho_{16}\rho_{35})],$$

$$(104)\quad \mu\begin{pmatrix}1\cdot\cdot\cdot\cdot\cdot\\ \cdot 1\cdot\cdot\cdot\cdot\\ \cdot\cdot\cdot 11\cdot\\ \cdot\cdot\cdot\cdot\cdot\end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^2\sigma_2\sigma_3\sigma_4^2\sigma_5\sigma_6[\rho_{12}(\rho_{15}\rho_{36}+\rho_{16}\rho_{35})+\rho_{13}(\rho_{15}\rho_{26}+\rho_{16}\rho_{25})$$
$$+(N-1)\rho_{12}\rho_{13}(\rho_{56}+\rho_{46}\rho_{45})+N\rho_{14}\{\rho_{15}(\rho_{26}\rho_{34}+\rho_{24}\rho_{36})$$
$$+\rho_{16}(\rho_{24}\rho_{35}+\rho_{25}\rho_{34})\}+\rho_{12}\rho_{14}(\rho_{35}\rho_{46}+\rho_{36}\rho_{45})+\rho_{13}\rho_{14}(\rho_{26}\rho_{45}+\rho_{25}\rho_{46})$$
$$+\rho_{14}^2(\rho_{25}\rho_{36}+\rho_{26}\rho_{35})+2\rho_{15}\rho_{16}\rho_{24}\rho_{34}$$
$$+(\rho_{13}\rho_{34}+\rho_{13}\rho_{24})(\rho_{15}\rho_{46}+\rho_{16}\rho_{45}+2\rho_{14}\rho_{56})],$$

$$(105)\quad \mu\begin{pmatrix} \cdot & 2 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} = \frac{N-1}{N^4}\sigma_1^2\sigma_2^2\sigma_3\sigma_4\sigma_5\sigma_6[(N-1)(1+\rho_{12}^2)(\rho_{35}\rho_{46}+\rho_{36}\rho_{45})$$
$$+2N(\rho_{13}\rho_{24}+\rho_{14}\rho_{23})(\rho_{15}\rho_{26}+\rho_{16}\rho_{25})+4(\rho_{13}\rho_{14}\rho_{25}\rho_{26}+\rho_{15}\rho_{16}\rho_{23}\rho_{24})$$
$$+2(\rho_{13}+\rho_{12}\rho_{23})(\rho_{15}\rho_{46}+\rho_{16}\rho_{45})+2(\rho_{14}+\rho_{12}\rho_{24})(\rho_{15}\rho_{36}+\rho_{16}\rho_{35})$$
$$+2(\rho_{23}+\rho_{12}\rho_{13})(\rho_{25}\rho_{46}+\rho_{26}\rho_{45})+2(\rho_{24}+\rho_{12}\rho_{14})(\rho_{25}\rho_{36}+\rho_{26}\rho_{35})],$$

$$(106)\quad \mu\begin{pmatrix} \cdot & 1 & 1 & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} = \frac{N-1}{N^4}\sigma_1^3\sigma_2\sigma_3\sigma_4\sigma_5\sigma_6[2\rho_{12}(\rho_{35}\rho_{46}+\rho_{36}\rho_{45})+2\rho_{13}(\rho_{25}\rho_{46}+\rho_{26}\rho_{45})$$
$$+2\rho_{14}(\rho_{25}\rho_{36}+\rho_{26}\rho_{35})+4\rho_{15}\rho_{16}(\rho_{12}\rho_{34}+\rho_{13}\rho_{24}+\rho_{14}\rho_{23})$$
$$+(N+1)\{(\rho_{23}+\rho_{12}\rho_{13})(\rho_{15}\rho_{46}+\rho_{16}\rho_{45})+(\rho_{24}+\rho_{12}\rho_{14})(\rho_{15}\rho_{36}+\rho_{16}\rho_{35})$$
$$+(\rho_{34}+\rho_{13}\rho_{14})(\rho_{15}\rho_{26}+\rho_{16}\rho_{25})\}],$$

$$(107)\quad \mu\begin{pmatrix} \cdot & 1 & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} = \frac{N-1}{N^4}\sigma_1^2\sigma_2^2\sigma_3\sigma_4\sigma_5\sigma_6[(1+\rho_{12}^2)(\rho_{35}\rho_{46}+\rho_{36}\rho_{45})$$
$$+(\rho_{13}+3\rho_{12}\rho_{23})(\rho_{15}\rho_{46}+\rho_{16}\rho_{45})+(\rho_{24}+3\rho_{12}\rho_{14})(\rho_{25}\rho_{36}+\rho_{26}\rho_{35})$$
$$+2\rho_{34}(\rho_{15}\rho_{16}+\rho_{25}\rho_{26})+2(\rho_{13}\rho_{14}\rho_{25}\rho_{26}+\rho_{15}\rho_{16}\rho_{23}\rho_{24})$$
$$+(\rho_{15}\rho_{26}+\rho_{16}\rho_{25})\{\rho_{13}\rho_{24}+(N+1)\rho_{12}\rho_{34}+(N+2)\rho_{14}\rho_{23}\}$$
$$+N\{(\rho_{23}+\rho_{12}\rho_{13})(\rho_{25}\rho_{46}+\rho_{26}\rho_{45})+(\rho_{14}+\rho_{12}\rho_{24})(\rho_{15}\rho_{36}+\rho_{16}\rho_{35})\}],$$

$$(108)\quad \mu\begin{pmatrix} \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} = \frac{N-1}{N^4}\sigma_1\sigma_2\sigma_3^2\sigma_4^2\sigma_5\sigma_6[(N-1)(\rho_{12}+\rho_{13}\rho_{23})(\rho_{56}+\rho_{45}\rho_{46})$$
$$+N\{(\rho_{14}\rho_{35}+\rho_{15}\rho_{34})(\rho_{24}\rho_{36}+\rho_{26}\rho_{34})+(\rho_{14}\rho_{36}+\rho_{16}\rho_{34})(\rho_{24}\rho_{35}+\rho_{25}\rho_{34})\}$$
$$+2\rho_{12}\rho_{34}\rho_{36}\rho_{45}+\rho_{13}\rho_{24}(\rho_{35}\rho_{46}+\rho_{36}\rho_{45})+\rho_{13}\rho_{26}(\rho_{35}+\rho_{34}\rho_{45})$$
$$+2\rho_{14}\rho_{34}(\rho_{25}\rho_{36}+\rho_{26}\rho_{35})+2\rho_{24}\rho_{34}(\rho_{15}\rho_{36}+\rho_{16}\rho_{35})+2\rho_{34}\rho_{36}(\rho_{12}\rho_{34}+\rho_{13}\rho_{24})$$
$$+(\rho_{15}+\rho_{14}\rho_{45})(\rho_{26}+\rho_{23}\rho_{36})+(\rho_{16}+\rho_{14}\rho_{46})(\rho_{25}+\rho_{23}\rho_{35})$$
$$+(\rho_{24}+\rho_{23}\rho_{34})(2\rho_{14}\rho_{56}+\rho_{15}\rho_{46}+\rho_{16}\rho_{45})+(\rho_{36}+\rho_{34}\rho_{46})(2\rho_{12}\rho_{35}+\rho_{13}\rho_{25})].$$

7 variates

$$(109)\quad \mu\begin{pmatrix} 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} = 2\frac{N-1}{N^4}\sigma_1^2\sigma_2\sigma_3\sigma_4\sigma_5\sigma_6\sigma_7[\rho_{12}\rho_{14}(\rho_{36}\rho_{57}+\rho_{37}\rho_{56})+\rho_{12}\rho_{15}(\rho_{36}\rho_{47}+\rho_{37}\rho_{46})$$
$$+\rho_{12}\rho_{16}(\rho_{34}\rho_{57}+\rho_{35}\rho_{47})+\rho_{12}\rho_{17}(\rho_{34}\rho_{56}+\rho_{35}\rho_{46})+\rho_{13}\rho_{14}(\rho_{26}\rho_{57}+\rho_{27}\rho_{56})$$
$$+\rho_{13}\rho_{15}(\rho_{26}\rho_{47}+\rho_{27}\rho_{46})+\rho_{13}\rho_{16}(\rho_{24}\rho_{57}+\rho_{25}\rho_{47})+\rho_{13}\rho_{17}(\rho_{24}\rho_{56}+\rho_{25}\rho_{46})$$
$$+\rho_{14}\rho_{16}(\rho_{25}\rho_{37}+\rho_{27}\rho_{35})+\rho_{14}\rho_{17}(\rho_{25}\rho_{36}+\rho_{26}\rho_{35})+\rho_{15}\rho_{16}(\rho_{24}\rho_{37}+\rho_{27}\rho_{34})$$
$$+\rho_{15}\rho_{17}(\rho_{24}\rho_{36}+\rho_{26}\rho_{34})+(N-1)\{\rho_{12}\rho_{13}(\rho_{46}\rho_{57}+\rho_{47}\rho_{56})$$
$$+\rho_{14}\rho_{15}(\rho_{26}\rho_{37}+\rho_{27}\rho_{36})+\rho_{16}\rho_{17}(\rho_{24}\rho_{35}+\rho_{25}\rho_{34})\}],$$

$$(110)\quad \mu\begin{pmatrix} \cdot & 1 & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} = \frac{N-1}{N^4}\sigma_1^2\sigma_2\sigma_3\sigma_4\sigma_5\sigma_6\sigma_7[(\rho_{24}+\rho_{12}\rho_{14})(\rho_{36}\rho_{57}+\rho_{37}\rho_{56})$$
$$+(\rho_{25}+\rho_{12}\rho_{15})(\rho_{36}\rho_{47}+\rho_{37}\rho_{46})+(\rho_{26}+\rho_{12}\rho_{16})(\rho_{34}\rho_{57}+\rho_{35}\rho_{47})$$
$$+(\rho_{27}+\rho_{12}\rho_{17})(\rho_{34}\rho_{56}+\rho_{35}\rho_{46})+(\rho_{13}\rho_{24}+\rho_{14}\rho_{23})(\rho_{16}\rho_{57}+\rho_{17}\rho_{56})$$
$$+(\rho_{13}\rho_{25}+\rho_{15}\rho_{23})(\rho_{16}\rho_{47}+\rho_{17}\rho_{46})+(\rho_{13}\rho_{26}+\rho_{16}\rho_{23})(\rho_{14}\rho_{57}+\rho_{15}\rho_{47})$$
$$+(\rho_{13}\rho_{27}+\rho_{17}\rho_{23})(\rho_{14}\rho_{56}+\rho_{15}\rho_{46})+(\rho_{14}\rho_{26}+\rho_{16}\rho_{24})(\rho_{15}\rho_{37}+\rho_{17}\rho_{35})$$
$$+(\rho_{14}\rho_{27}+\rho_{17}\rho_{24})(\rho_{15}\rho_{36}+\rho_{16}\rho_{35})+(\rho_{15}\rho_{26}+\rho_{16}\rho_{25})(\rho_{14}\rho_{37}+\rho_{17}\rho_{34})$$
$$+(\rho_{15}\rho_{27}+\rho_{17}\rho_{25})(\rho_{14}\rho_{36}+\rho_{16}\rho_{34})+(N-1)\{(\rho_{23}+\rho_{12}\rho_{13})(\rho_{46}\rho_{57}+\rho_{47}\rho_{56})$$
$$+(\rho_{14}\rho_{25}+\rho_{15}\rho_{24})(\rho_{16}\rho_{37}+\rho_{17}\rho_{36})+(\rho_{14}\rho_{35}+\rho_{15}\rho_{34})(\rho_{16}\rho_{27}+\rho_{17}\rho_{26})\}].$$

8 *variates*

$$(111)\ \mu\begin{pmatrix}\cdot 1\cdot\cdot\cdot\cdot\cdot\cdot\\ \cdot\cdot\cdot 1\cdot\cdot\cdot\cdot\\ \cdot\cdot 1\cdot\cdot\cdot\cdot\cdot\\ \cdot\cdot\cdot\cdot 1\cdot\cdot\cdot\\ \cdot\cdot\cdot 1\cdot\cdot\cdot\\ \cdot\cdot\cdot\cdot\cdot 1\cdot\\ \cdot\cdot\cdot\cdot\cdot\cdot 1\end{pmatrix} = \frac{N-1}{N^4}\sigma_1\sigma_2\sigma_3\sigma_4\sigma_5\sigma_6\sigma_7\sigma_8[(\rho_{13}\rho_{25}+\rho_{15}\rho_{23})(\rho_{47}\rho_{68}+\rho_{48}\rho_{67})$$
$$+(\rho_{13}\rho_{26}+\rho_{16}\rho_{23})(\rho_{47}\rho_{58}+\rho_{48}\rho_{57})+(\rho_{13}\rho_{27}+\rho_{17}\rho_{23})(\rho_{45}\rho_{68}+\rho_{46}\rho_{58})$$
$$+(\rho_{13}\rho_{28}+\rho_{18}\rho_{23})(\rho_{45}\rho_{67}+\rho_{46}\rho_{57})+(\rho_{14}\rho_{25}+\rho_{15}\rho_{24})(\rho_{37}\rho_{68}+\rho_{38}\rho_{67})$$
$$+(\rho_{14}\rho_{26}+\rho_{16}\rho_{24})(\rho_{37}\rho_{58}+\rho_{38}\rho_{57})+(\rho_{14}\rho_{27}+\rho_{17}\rho_{24})(\rho_{35}\rho_{68}+\rho_{36}\rho_{58})$$
$$+(\rho_{14}\rho_{28}+\rho_{18}\rho_{24})(\rho_{35}\rho_{67}+\rho_{36}\rho_{57})+(\rho_{15}\rho_{27}+\rho_{17}\rho_{25})(\rho_{36}\rho_{48}+\rho_{38}\rho_{46})$$
$$+(\rho_{15}\rho_{28}+\rho_{18}\rho_{25})(\rho_{36}\rho_{47}+\rho_{37}\rho_{46})+(\rho_{16}\rho_{27}+\rho_{17}\rho_{26})(\rho_{35}\rho_{48}+\rho_{38}\rho_{45})$$
$$+(\rho_{16}\rho_{28}+\rho_{18}\rho_{26})(\rho_{35}\rho_{47}+\rho_{37}\rho_{45})+(N-1)\{(\rho_{13}\rho_{24}+\rho_{14}\rho_{23})(\rho_{57}\rho_{68}+\rho_{58}\rho_{67})$$
$$+(\rho_{15}\rho_{26}+\rho_{16}\rho_{25})(\rho_{37}\rho_{48}+\rho_{38}\rho_{47})+(\rho_{17}\rho_{28}+\rho_{18}\rho_{27})(\rho_{35}\rho_{46}+\rho_{36}\rho_{45})\}].$$

FREQUENCY DISTRIBUTION OF AN INDEX WHEN BOTH THE COMPONENTS FOLLOW THE NORMAL LAW.

By A. S. MERRILL.

When two sets of related measurements are made, each of which follows the normal law, it is usually supposed that the index, the quotient of a value from one set divided by the corresponding value from the other set, will also be practically normally distributed. Thus cephalic indices in skull measurements are stated to follow the normal law very closely. In the sequel the conditions are shown under which an index distribution on the hypothesis of the normal distribution of components will itself be close to the normal, and in an approximate way, by graphs, the degree to which an index distribution will differ from normality in cases likely to arise in practice.

Suppose y denotes the variate in one set as measured from its mean $\bar{y}$; and that x is the second variate measured from its mean $\bar{x}$. Then the raw measures are $\bar{y}+y$ and $\bar{x}+x$. Assuming that r is the coefficient of correlation between x and y, then the equation of the correlation surface is

$$z = z_0 e^{-\frac{1}{2}\frac{1}{1-r^2}\left(\frac{x^2}{\sigma_1^2} - \frac{2rxy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right)},$$

σ_1 and σ_2 being the standard deviations of x and y respectively.

The index
$$I = \frac{\bar{y}+y}{\bar{x}+x}$$
may be expressed as a series, for
$$I = \frac{\bar{y}}{\bar{x}}\left(1+\frac{y}{\bar{y}}\right)\left(1+\frac{x}{\bar{x}}\right)^{-1}.$$

Designating $\frac{\bar{y}}{\bar{x}}$ by I_0, we have

$$I = I_0\left\{1 + \frac{y}{\bar{y}} - \frac{x}{\bar{x}} - \frac{xy}{\bar{x}\bar{y}} + \frac{x^2}{\bar{x}^2} + \frac{x^2y}{\bar{x}^2\bar{y}} - \frac{x^3}{\bar{x}^3} - \frac{x^3y}{\bar{x}^3\bar{y}} + \frac{x^4}{\bar{x}^4} + \frac{x^4y}{\bar{x}^4\bar{y}} - \frac{x^5}{\bar{x}^5} - \frac{x^5y}{\bar{x}^5y} + \frac{x^6}{\bar{x}^6} + \frac{x^6y}{\bar{x}^6\bar{y}} - \frac{x^7}{\bar{x}^7} - \frac{x^7y}{\bar{x}^7\bar{y}} + \frac{x^8}{\bar{x}^8}\right\}.$$

Summing and dividing by their number* we obtain the mean value of I:

$$\bar{I} = I_0\{1 + v_1(v_1 - rv_2)(1 + 3v_1^2 + 15v_1^4 + 105v_1^6)\} = \nu_1,$$

* For the mean products used here and later see K. Pearson and A. W. Young, "On the Product Moments of Various Orders of the Normal Correlation Surface of Two Variates," *Biometrika*, Vol. XII. (1918—19), pp. 86 ff.

where v_1 and v_2 are one one-hundredth of the coefficients of variation* of x and y respectively, and ν_i is used to denote the i-th moment about the original origin. It may be noted that in the process of summing for this and the following moments, the terms of odd degree on our hypothesis vanish. We shall drop all terms of degree above the eighth. The terms of degree nine will therefore contribute nothing, and will accordingly be omitted entirely in the expressions for the various powers of I. The value of v_1 or v_2 is usually rather less than ·1. In omitting a v_1^9 or v_2^9, we are omitting a quantity of the order $(\cdot 1)^9$, sufficiently small to be negligible in most cases.

By multiplying we obtain the following results:

$$I^2 = I_0^2\Big\{1 - \frac{2x}{\bar{x}} + \frac{2y}{\bar{y}} + \frac{3x^2}{\bar{x}^2} - \frac{4xy}{\bar{x}\bar{y}} + \frac{y^2}{\bar{y}^2} - \frac{4x^3}{\bar{x}^3} + \frac{6x^2y}{\bar{x}^2\bar{y}} - \frac{2xy^2}{\bar{x}\bar{y}^2}$$
$$+ \frac{5x^4}{\bar{x}^4} - \frac{8x^3y}{\bar{x}^3\bar{y}} + \frac{3x^2y^2}{\bar{x}^2\bar{y}^2} - \frac{6x^5}{\bar{x}^5} + \frac{10x^4y}{\bar{x}^4\bar{y}} - \frac{4x^3y^2}{\bar{x}^3\bar{y}^2} + \frac{7x^6}{\bar{x}^6}$$
$$- \frac{12x^5y}{\bar{x}^5\bar{y}} + \frac{5x^4y^2}{\bar{x}^4\bar{y}^2} - \frac{8x^7}{\bar{x}^7} + \frac{14x^6y}{\bar{x}^6\bar{y}} - \frac{6x^5y^2}{\bar{x}^5\bar{y}^2} + \frac{9x^8}{\bar{x}^8} - \frac{16x^7y}{\bar{x}^7\bar{y}} + \frac{7x^6y^2}{\bar{x}^6\bar{y}^2}\Big\}.$$

$$I^3 = I_0^3\Big\{1 - \frac{3x}{\bar{x}} + \frac{3y}{\bar{y}} + \frac{6x^2}{\bar{x}^2} - \frac{9xy}{\bar{x}\bar{y}} + \frac{3y^2}{\bar{y}^2} - \frac{10x^3}{\bar{x}^3} + \frac{18x^2y}{\bar{x}^2\bar{y}} - \frac{9xy^2}{\bar{x}\bar{y}^2}$$
$$+ \frac{y^3}{\bar{y}^3} + \frac{15x^4}{\bar{x}^4} - \frac{30x^3y}{\bar{x}^3\bar{y}} + \frac{18x^2y^2}{\bar{x}^2\bar{y}^2} - \frac{3xy^3}{\bar{x}\bar{y}^3} - \frac{21x^5}{\bar{x}^5} + \frac{45x^4y}{\bar{x}^4\bar{y}}$$
$$- \frac{30x^3y^2}{\bar{x}^3\bar{y}^2} + \frac{6x^2y^3}{\bar{x}^2\bar{y}^3} + \frac{28x^6}{\bar{x}^6} - \frac{63x^5y}{\bar{x}^5\bar{y}} + \frac{45x^4y^2}{\bar{x}^4\bar{y}^2} - \frac{10x^3y^3}{\bar{x}^3\bar{y}^3} - \frac{36x^7}{\bar{x}^7}$$
$$+ \frac{84x^6y}{\bar{x}^6\bar{y}} - \frac{63x^5y^2}{\bar{x}^5\bar{y}^2} + \frac{15x^4y^3}{\bar{x}^4\bar{y}^3} + \frac{45x^8}{\bar{x}^8} - \frac{108x^7y}{\bar{x}^7\bar{y}} + \frac{84x^6y^2}{\bar{x}^6\bar{y}^2} - \frac{21x^5y^3}{\bar{x}^5\bar{y}^3}\Big\}.$$

$$I^4 = I_0^4\Big\{1 - \frac{4x}{\bar{x}} + \frac{4y}{\bar{y}} + \frac{10x^2}{\bar{x}^2} - \frac{16xy}{\bar{x}\bar{y}} + \frac{6y^2}{\bar{y}^2} - \frac{20x^3}{\bar{x}^3} + \frac{40x^2y}{\bar{x}^2\bar{y}}$$
$$- \frac{24xy^2}{\bar{x}\bar{y}^2} + \frac{4y^3}{\bar{y}^3} + \frac{35x^4}{\bar{x}^4} - \frac{80x^3y}{\bar{x}^3\bar{y}} + \frac{60x^2y^2}{\bar{x}^2\bar{y}^2} - \frac{16xy^3}{\bar{x}\bar{y}^3}$$
$$+ \frac{y^4}{\bar{y}^4} - \frac{56x^5}{\bar{x}^5} + \frac{140x^4y}{\bar{x}^4\bar{y}} - \frac{120x^3y^2}{\bar{x}^3\bar{y}^2} + \frac{40x^2y^3}{\bar{x}^2\bar{y}^3}$$
$$- \frac{4xy^4}{\bar{x}\bar{y}^4} + \frac{84x^6}{\bar{x}^6} - \frac{224x^5y}{\bar{x}^5\bar{y}} + \frac{210x^4y^2}{\bar{x}^4\bar{y}^2} - \frac{80x^3y^3}{\bar{x}^3\bar{y}^3} + \frac{10x^2y^4}{\bar{x}^2\bar{y}^4}$$
$$- \frac{120x^7}{\bar{x}^7} + \frac{336x^6y}{\bar{x}^6\bar{y}} - \frac{336x^5y^2}{\bar{x}^5\bar{y}^2} + \frac{140x^4y^3}{\bar{x}^4\bar{y}^3} - \frac{20x^3y^4}{\bar{x}^3\bar{y}^4}$$
$$+ \frac{165x^8}{\bar{x}^8} - \frac{480x^7y}{\bar{x}^7\bar{y}} + \frac{504x^6y^2}{\bar{x}^6\bar{y}^2} - \frac{224x^5y^3}{\bar{x}^5\bar{y}^3} + \frac{35x^4y^4}{\bar{x}^4\bar{y}^4}\Big\}.$$

* The coefficient of variation is usually defined to be the *per cent.* of the standard deviation on the mean.

$$I^5 = I_0^5\Big\{1 - \frac{5x}{\bar{x}} + \frac{5y}{\bar{y}} + \frac{15x^2}{\bar{x}^2} - \frac{25xy}{\bar{x}\bar{y}} + \frac{10y^2}{\bar{y}^2} - \frac{35x^3}{\bar{x}^3} + \frac{75x^2y}{\bar{x}^2\bar{y}}$$
$$- \frac{50xy^2}{\bar{x}\bar{y}^2} + \frac{10y^3}{\bar{y}^3} + \frac{70x^4}{\bar{x}^4} - \frac{175x^3y}{\bar{x}^3\bar{y}} + \frac{150x^2y^2}{\bar{x}^2\bar{y}^2} - \frac{50xy^3}{\bar{x}\bar{y}^3} + \frac{5y^4}{\bar{y}^4}$$
$$- \frac{126x^5}{\bar{x}^5} + \frac{350x^4y}{\bar{x}^4\bar{y}} - \frac{350x^3y^2}{\bar{x}^3\bar{y}^2} + \frac{150x^2y^3}{\bar{x}^2\bar{y}^3} - \frac{25xy^4}{\bar{x}\bar{y}^4} + \frac{y^5}{\bar{y}^5} + \frac{210x^6}{\bar{x}^6}$$
$$- \frac{630x^5y}{\bar{x}^5\bar{y}} + \frac{700x^4y^2}{\bar{x}^4\bar{y}^2} - \frac{350x^3y^3}{\bar{x}^3\bar{y}^3} + \frac{75x^2y^4}{\bar{x}^2\bar{y}^4} - \frac{5xy^5}{\bar{x}\bar{y}^5}$$
$$- \frac{330x^7}{\bar{x}^7} + \frac{1050x^6y}{\bar{x}^6\bar{y}} - \frac{1260x^5y^2}{\bar{x}^5\bar{y}^2} + \frac{700x^4y^3}{\bar{x}^4\bar{y}^3} - \frac{175x^3y^4}{\bar{x}^3\bar{y}^4} + \frac{15x^2y^5}{\bar{x}^2\bar{y}^5}$$
$$+ \frac{495x^8}{\bar{x}^8} - \frac{1650x^7y}{\bar{x}^7\bar{y}} + \frac{2100x^6y^2}{\bar{x}^6\bar{y}^2} - \frac{1260x^5y^3}{\bar{x}^5\bar{y}^3} + \frac{350x^4y^4}{\bar{x}^4\bar{y}^4} - \frac{35x^3y^5}{\bar{x}^3\bar{y}^5}\Big\}.$$

Summing these results and taking mean values we obtain by aid of the higher product moments of the normal surface the higher moment coefficients of the index about the original axes:

$$\nu_2' = I_0^2\{1 + 3v_1^2 - 4rv_1v_2 + v_2^2 + 15v_1^4 - 24rv_1^3v_2 + 3v_1^2v_2^2$$
$$+ 6r^2v_1^2v_2^2 + 105v_1^6 - 180rv_1^5v_2 + 15v_1^4v_2^2 + 60r^2v_1^4v_2^2$$
$$+ 945v_1^8 - 1680rv_1^7v_2 + 105v_1^6v_2^2 + 630r^2v_1^6v_2^2\},$$

$$\nu_3' = I_0^3\{1 + 6v_1^2 - 9rv_1v_2 + 3v_2^2 + 45v_1^4 - 90rv_1^3v_2 + 18v_1^2v_2^2$$
$$+ 36r^2v_1^2v_2^2 - 9rv_1v_2^3 + 420v_1^6 - 945rv_1^5v_2 + 135v_1^4v_2^2$$
$$+ 540r^2v_1^4v_2^2 - 90rv_1^3v_2^3 - 60r^3v_1^3v_2^3 + 4725v_1^8 - 11340rv_1^7v_2$$
$$+ 1260v_1^6v_2^2 + 7560r^2v_1^6v_2^2 - 945rv_1^5v_2^3 - 1260r^3v_1^5v_2^3\},$$

$$\nu_4' = I_0^4\{1 + 10v_1^2 - 16rv_1v_2 + 6v_2^2 + 105v_1^4 - 240rv_1^3v_2$$
$$+ 60v_1^2v_2^2 + 120r^2v_1^2v_2^2 - 48rv_1v_2^3 + 3v_2^4 + 1260v_1^6$$
$$- 3360rv_1^5v_2 + 630v_1^4v_2^2 + 2520r^2v_1^4v_2^2$$
$$- 720rv_1^3v_2^2 - 480r^3v_1^3v_2^3 + 30v_1^2v_2^4 + 120r^2v_1^2v_2^4$$
$$+ 17325v_1^8 - 50400rv_1^7v_2 + 7560v_1^6v_2^2 + 45360r^2v_1^6v_2^2$$
$$- 10080rv_1^5v_2^3 - 13440r^3v_1^5v_2^3 + 315v_1^4v_2^4$$
$$+ 2520r^2v_1^4v_2^4 + 840r^4v_1^4v_2^4\},$$

$$\nu_5' = I_0^5\{1 + 15v_1^2 - 25rv_1v_2 + 10v_2^2 + 210v_1^4 - 525rv_1^3v_2$$
$$+ 150v_1^2v_2^2 + 300r^2v_1^2v_2^2 - 150rv_1v_2^3 + 15v_2^4 + 3150v_1^6$$
$$- 9450rv_1^5v_2 + 2100v_1^4v_2^2 + 8400r^2v_1^4v_2^2 - 3150rv_1^3v_2^3$$
$$- 2100r^3v_1^3v_2^3 + 225v_1^2v_2^4 + 900r^2v_1^2v_2^4 - 75rv_1v_2^5$$
$$+ 51975v_1^8 - 173250rv_1^7v_2 + 31500v_1^6v_2^2 + 189000r^2v_1^6v_2^2$$
$$- 56700rv_1^5v_2^3 - 75600r^3v_1^5v_2^3 + 3150v_1^4v_2^4 + 25200r^2v_1^4v_2^4$$
$$+ 8400r^4v_1^4v_2^4 - 1575rv_1^3v_2^5 - 2100r^3v_1^3v_2^5\}.$$

It is needful to take the I's to a high power in the v's, if we desire to obtain the β's to a reasonable degree of approximation.

Transferring the origin to the mean of I we have for the first four moment coefficients:

$$\begin{aligned}
\mu_2 &= \nu_2' - \nu_1'^2 \\
&= I_0^2 \{v_1^2 - 2rv_1v_2 + v_2^2 + 8v_1^4 - 16rv_1^3v_2 + 3v_1^2v_2^2 \\
&\qquad + 5r^2v_1^2v_2^2 + 69v_1^6 - 138rv_1^5v_2 + 15v_1^4v_2^2 + 54r^2v_1^4v_2^2 \\
&\qquad + 696v_1^8 - 1392rv_1^7v_2 + 105v_1^6v_2^2 + 591r^2v_1^6v_2^2\}, \\
\mu_3 &= \nu_3' - 3\nu_1'\nu_2' + 2\nu_1'^3 \\
&= I_0^3 \{6v_1^4 - 18rv_1^3v_2 + 6v_1^2v_2^2 + 12r^2v_1^2v_2^2 - 6rv_1v_2^3 \\
&\qquad + 116v_1^6 - 348rv_1^5v_2 + 72v_1^4v_2^2 + 276r^2v_1^4v_2^2 \\
&\qquad - 72rv_1^3v_2^3 - 44r^2v_1^3v_2^3 + 1872v_1^8 - 5616rv_1^7v_2 \\
&\qquad + 828v_1^6v_2^2 + 4788r^2v_1^6v_2^2 - 828rv_1^5v_2^3 - 1044r^2v_1^5v_2^3\}, \\
\mu_4 &= \nu_4' - 4\nu_1'\nu_3' + 6\nu_1'^2\nu_2' - 3\nu_1'^4 \\
&= I_0^4 \{3v_1^4 - 12rv_1^3v_2 + 6v_1^2v_2^2 + 12r^2v_1^2v_2^2 - 12rv_1v_2^3 + 3v_2^4 \\
&\qquad + 120v_1^6 - 480rv_1^5v_2 + 150v_1^4v_2^2 + 570r^2v_1^4v_2^2 \\
&\qquad - 300rv_1^3v_2^3 - 180r^2v_1^3v_2^3 + 30v_1^2v_2^4 + 90r^2v_1^2v_2^4 \\
&\qquad + 3054v_1^8 - 12216rv_1^7v_2 + 15588r^2v_1^6v_2^2 + 2736v_1^6v_2^2 \\
&\qquad - 6744r^2v_1^5v_2^3 - 5472rv_1^5v_2^3 + 2106r^2v_1^4v_2^4 \\
&\qquad + 633r^4v_1^4v_2^4 + 315v_1^4v_2^4\}, \\
\mu_5 &= \nu_5' - 5\nu_1'\nu_4' + 10\nu_1'^2\nu_3' - 10\nu_1'^3\nu_2' + 4\nu_1'^5 \\
&= I_0^5 \{60v_1^6 - 300rv_1^5v_2 + 120v_1^4v_2^2 + 480r^2v_1^4v_2^2 - 360rv_1^3v_2^3 \\
&\qquad - 240r^3v_1^3v_2^3 + 60v_1^2v_2^4 + 240r^2v_1^2v_2^4 - 60rv_1v_2^5 \\
&\qquad + 2960v_1^8 - 14800rv_1^7v_2 + 25260r^2v_1^6v_2^2 + 4340v_1^6v_2^2 \\
&\qquad - 16580r^3v_1^5v_2^3 - 13020rv_1^5v_2^3 + 3160r^4v_1^4v_2^4 \\
&\qquad + 10260r^2v_1^4v_2^4 + 1380v_1^4v_2^4 - 1380rv_1^3v_2^5 - 1580r^3v_1^3v_2^5\}.
\end{aligned}$$

These expressions assume a more compact form by the introduction of two auxiliary quantities, p_2 and λ, defined by the following equations:

$$p_2 = v_1^2 - 2rv_1v_2 + v_2^2,$$

$$\lambda = \frac{v_1 - rv_2}{\sqrt{v_1^2 - 2rv_1v_2 + v_2^2}} = \frac{v_1 - rv_2}{\sqrt{p_2}};$$

thus the sign of $\lambda\sqrt{p_2}$ is defined by that of $v_1 - rv_2$.

Since $$v_1^2 - 2rv_1v_2 + v_2^2 = (v_1 - rv_2)^2 + v_2^2(1 - r^2),$$

$$p_2 = \lambda^2 p_2 + v_2^2(1 - r^2),$$

hence $$v_2^2(1 - r^2) = p_2(1 - \lambda^2).$$

The expression p_2 is of the second order in v_1 or v_2; λ is of order zero.

Substituting we have

$$\begin{aligned}
\mu_2 &= I_0^2 p_2 \{1 + v_1^2(5\lambda^2 + 3) + v_1^4(54\lambda^2 + 15) + v_1^6(591\lambda^2 + 105)\}, \\
\mu_3 &= I_0^3 p_2^{\frac{3}{2}} v_1\lambda \{6 + 4v_1^2(11\lambda^2 + 18) + v_1^4(1044\lambda^2 + 828)\}, \\
\mu_4 &= I_0^4 p_2^2 \{3 + 30v_1^2(3\lambda^2 + 1) + v_1^4(633\lambda^4 + 2106\lambda^2 + 315)\}, \\
\mu_5 &= I_0^5 p_2^{\frac{5}{2}} v_1\lambda \{60 + v_1^2(1580\lambda^2 + 1380)\}.
\end{aligned}$$

Accordingly expanding in powers of v_1:

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = v_1^2\lambda^2\{36 - v_1^2(12\lambda^2 - 540) + 4v_1^4(-146\lambda^4 + 450\lambda^2 + 1917)\},$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = 3 + 3\{v_1^2(20\lambda^2 + 4) + v_1^4(-14\lambda^4 + 404\lambda^2 + 42)\},$$

$$\beta_3' = \frac{\mu_5}{\mu_2^{\frac{5}{2}}} = 10v_1\lambda\{6 + v_1^2(83\lambda^2 + 93)\}.$$

These functions are thus dependent only upon v_1 and λ. As mentioned above, v_1 is a coefficient of variation. Its value will usually lie between 0·01 and 0·1, and in the case of anthropometric measurements (weight, of course, excluded) is found most commonly between 0·02 and 0·04. The expression for λ shows that this function of v_1, v_2 and r is always less than unity in numerical value. Its nature can readily be seen by considering the distribution of the I if approximations to the *first order only* in v_1 or v_2 are considered. In this case

$$I = I_0\left(1 + \frac{y}{\bar{y}} - \frac{x}{\bar{x}}\right) \text{ and } \bar{I} = I_0.$$

Then

$$y = \bar{y}\left(\frac{I}{\bar{I}} - 1 + \frac{x}{\bar{x}}\right).$$

Substituting this value for y in the equation of the normal correlation surface, we have

$$z = z_0 e^{-\frac{1}{2}\frac{p_2}{v_2^2(1-r^2)}\left\{\frac{x^2}{\sigma_1^2} - \frac{2(rv_2 - v_1)\left(\frac{I}{\bar{I}} - 1\right)x}{\sigma_1 p_2} + \frac{\left(\frac{I}{\bar{I}} - 1\right)^2}{p_2}\right\}}.$$

Let

$$\omega = \frac{I}{\bar{I}} - 1 = \frac{y}{\bar{y}} - \frac{x}{\bar{x}}.$$

Then

$$\omega^2 = \frac{y^2}{\bar{y}^2} - \frac{2xy}{\bar{x}\bar{y}} + \frac{x^2}{\bar{x}^2},$$

and

$$\sigma_\omega = \sqrt{v_2^2 - 2rv_1v_2 + v_1^2} = \sqrt{p_2}.$$

Thus we have as the correlation surface

$$z = z_0 e^{-\frac{1}{2}\frac{1}{1-\lambda^2}\left(\frac{x^2}{\sigma_1^2} + \frac{2\lambda x\omega}{\sigma_1\sigma_\omega} + \frac{\omega^2}{\sigma_\omega^2}\right)},$$

showing that λ is, with reversed sign, the coefficient of correlation of x and ω.

In Table I are given the values of β_1, β_2 and β_3' for $\lambda^2 = 0$, ·05, ·1, ·2, ·3, ·4, ·5, ·6, ·7, ·8, ·9, 1, and for $v_1 =$ ·01, ·04, ·07, 0·1. The signs are given as if λ were positive, but since a change of sign of λ only changes the signs of μ_2 and μ_5, the table suffices also for negative values of λ. The value of λ^2 will seldom be as great as ·5, which corresponds to a correlation coefficient (of x and ω as shown above) of over ·70. Coefficients of variation greater than about ·04 are also unusual. In the portion of the table covering values likely to occur, it may be observed that β_1 is always small and β_2 differs but little from 3. For the larger values of both v_1 and λ^2, this

TABLE I. *Values of β_1, β_2 and β_3' of the Frequency Distribution of Index.*

v_1	λ^2	0·00	0·05	0·10	0·20	0·30	0·40	0·50	0·60	0·70	0·80	0·90	1·00
·01	β_1	0·0000	0·0002	0·0004	0·0007	0·0011	0·0014	0·0018	0·0022	0·0025+	0·0029	0·0032	0·0036
	β_2	3·0012	3·0015+	3·0018	3·0024	3·0030	3·0036	3·0042	3·0048	3·0054	3·0060	3·0066	3·0072
	β_3'	0·0000	0·1344	0·1900	0·2688	0·3293	0·3803	0·4252	0·4659	0·5033	0·5381	0·5708	0·6018
·04	β_1	0·0000	0·0030	0·0059	0·0118	0·0177	0·0236	0·0295+	0·0354	0·0413	0·0472	0·0531	0·0590
	β_2	3·0195+	3·0245−	3·0294	3·0393	3·0492	3·0591	3·0690	3·0789	3·0888	3·0987	3·1086	3·1185+
	β_3'	0·0000	0·5505+	0·7794	1·1047	1·3559	1·5690	1·7579	1·9298	2·0889	2·2379	2·3787	2·5126
·07	β_1	0·0000	0·0095+	0·0190	0·0380	0·0571	0·0761	0·0951	0·1141	0·1331	0·1521	0·1711	0·1901
	β_2	3·0618	3·0780	3·0941	3·1264	3·1587	3·1909	3·2231	3·2553	3·2875+	3·3197	3·3518	3·3839
	β_3'	0·0000	1·0136	1·4380	2·0464	2·5219	2·9301	3·2961	3·6327	3·9476	4·2456	4·5302	4·8037
·10	β_1	0·0000	0·0211	0·0422	0·0844	0·1267	0·1687	0·2109	0·2531	0·2953	0·3374	0·3796	0·4217
	β_2	3·1326	3·1686	3·2047	3·2767	3·3486	3·4204	3·4922	3·5638	3·6354	3·7069	3·7783	3·8496
	β_3'	0·0000	1·5588	2·2177	3·1734	3·9321	4·5929	5·1937	5·7537	6·2837	6·7923	7·2830	7·7600

TABLE II. *Coefficients of τ_4, τ_5 and τ_6, respectively, in the Tetrachoric Series for the Frequency Distribution of Index.*

v_1	λ^2	0·00	0·05	0·10	0·20	0·30	0·40	0·50	0·60	0·70	0·80	0·90	1·00
·01	τ_4	0·0000	0·0110	0·0155+	0·0219	0·0269	0·0310	0·0347	0·0380	0·0410	0·0439	0·0465+	0·0490
	τ_5	0·0005+	0·0007	0·0008	0·0011	0·0014	0·0016	0·0019	0·0022	0·0025−	0·0027	0·0030	0·0033
	τ_6	0·0000	0·0000	0·0000	0·0001	0·0001	0·0001	0·0001	0·0002	0·0002	0·0002	0·0003	0·0003
·04	τ_4	0·0000	0·0444	0·0627	0·0887	0·1086	0·1254	0·1402	0·1536	0·1659	0·1774	0·1881	0·1983
	τ_5	0·0089	0·0112	0·0134	0·0179	0·0224	0·0270	0·0315−	0·0360	0·0405	0·0451	0·0496	0·0541
	τ_6	0·0000	0·0016	0·0025+	0·0041	0·0057	0·0073	0·0090	0·0108	0·0127	0·0146	0·0167	0·0188
·07	τ_4	0·0000	0·0796	0·1126	0·1593	0·1951	0·2252	0·2518	0·2758	0·2979	0·3185−	0·3378	0·3560
	τ_5	0·0282	0·0356	0·0430	0·0577	0·0724	0·0871	0·1018	0·1165+	0·1312	0·1459	0·1606	0·1752
	τ_6	0·0000	0·0085+	0·0131	0·0214	0·0297	0·0384	0·0474	0·0569	0·0669	0·0772	0·0880	0·0991
·10	τ_4	0·0000	0·1186	0·1677	0·2371	0·2907	0·3354	0·3750−	0·4108	0·4437	0·4743	0·5030	0·5302
	τ_5	0·0605+	0·0770	0·0934	0·1263	0·1591	0·1919	0·2247	0·2573	0·2900	0·3227	0·3552	0·3878
	τ_6	0·0000	0·0239	0·0367	0·0602	0·0832	0·1085+	0·1344	0·1616	0·1900	0·2199	0·2509	0·2832

is not the case. The value of β_3' shows a rather greater deviation from zero than might perhaps have been expected. In the smaller values, however, it is not sufficiently large to make the distribution differ to any considerable degree from normal.

This table can be used in the following manner in practical work. Values of v_1, v_2 and r having been found in two sets of measurements, λ^2 can be readily computed. From this value and that of v_1, the approximate values of β_1, β_2 and β_3' can be obtained to a fair degree of accuracy by interpolation.

The late Professor Thiele was, I believe, the first to suggest an expansion for frequency distributions in a series of differentials of the ordinates of the normal curve. Professor Pearson has given* an expansion in tetrachoric functions for any function, in the form

$$\frac{z}{N/\sigma} = \tau_1 + \frac{4}{\sqrt{24}}\sqrt{\beta_1}\,\tau_4 + \frac{5}{\sqrt{120}}(\beta_2 - 3)\,\tau_5 + \frac{6}{\sqrt{720}}(\beta_3' - 10\sqrt{\beta_1})\,\tau_6 + \ldots,$$

in which the tetrachoric functions differ from Thiele's functions only by a numerical factor. Thus if I is the variable, the argument is $(I - \bar{I})/\sigma$, where σ is the standard deviation of I and τ_s is defined by the equation

$$\tau_s\left(\frac{I-\bar{I}}{\sigma}\right) = \frac{1}{\sqrt{s\,!}}\left(-\frac{d}{d\left(\frac{I-\bar{I}}{\sigma}\right)}\right)^{s-1}\frac{1}{\sqrt{2\pi}\,\sigma}\,e^{-\frac{1}{2}\left(\frac{(I-\bar{I})}{\sigma^2}\right)^2}.$$

Table II has accordingly been formed showing the values of

$$\frac{4}{\sqrt{24}}\sqrt{\beta_1}, \quad \frac{5}{\sqrt{120}}(\beta_2 - 3) \text{ and } \frac{6}{\sqrt{720}}(\beta_3' - 10\sqrt{\beta_1})$$

for the above listed values of λ^2 and v_1. This table shows that the values of β_3' in the more useful part of the table will not shift the distribution much from normal, a fact which was not apparent from mere inspection of Table I. It further shows that for higher values of λ^2 and v_1, the fourth coefficient in the tetrachoric expansion is commensurable with the third. Since τ_6 is, in general†, about 4/5 as great as τ_5, a not wholly satisfactory result, to be shown graphically in the sequel, may be anticipated for such values.

For constructing graphs to illustrate possible cases, the procedure was as follows: In Table II above the values of the coefficients of the τ's of the tetrachoric expansion were found for the values of λ^2 and v_1 to be used. These were then multiplied by the values of the respective τ's as given in Dr Lee's tables for the values of h selected. The argument h is measured with the standard deviation as a unit. This process gave values of

$$y_1 = \frac{4}{\sqrt{24}}\sqrt{\beta_1}\tau_4,$$

$$y_2 = \frac{5}{\sqrt{120}}(\beta_2 - 3)\,\tau_5,$$

and

$$y_3 = \frac{6}{\sqrt{720}}(\beta_3' - 10\sqrt{\beta_1})\,\tau_6.$$

* "The Fifteen Constant Bivariate Frequency Surface," *Biometrika*, Vol. XVII. (1925), pp. 268 ff.

† See Dr Alice Lee's "Table of Tetrachoric Functions," *Biometrika*, Vol. XVII. (1925), pp. 351—354.

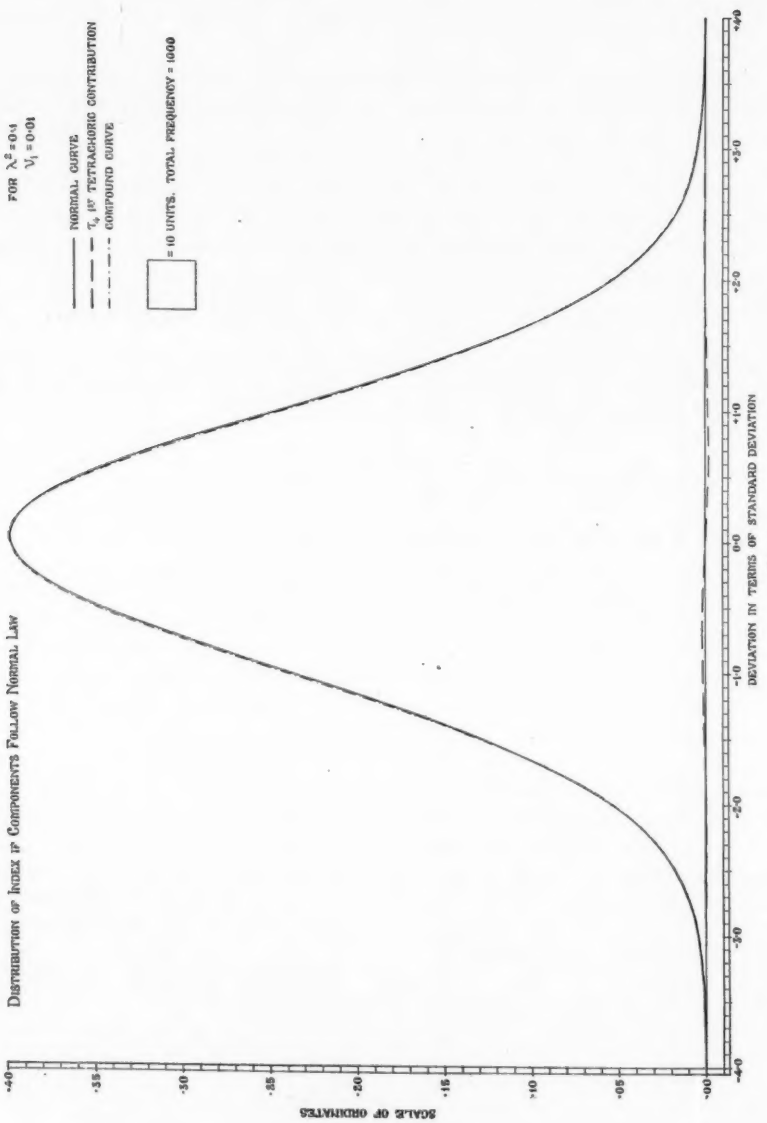

DISTRIBUTION OF INDEX IF COMPONENTS FOLLOW NORMAL LAW
FOR $\lambda^2 = 0{\cdot}1$
$V_1 = 0{\cdot}01$
NORMAL CURVE
$\frac{1}{\sqrt{2}}$ TETRACHORIC CONTRIBUTION
COMPOUND CURVE
= 10 UNITS. TOTAL FREQUENCY = 1000
SCALE OF ORDINATES
·40
·35
·30
·25
·20
·15
·10
·05
·00
-4·0
-3·0
-2·0
-1·0
0·0
+1·0
+2·0
+3·0
+4·0
DEVIATION IN TERMS OF STANDARD DEVIATION

Fig. 1.

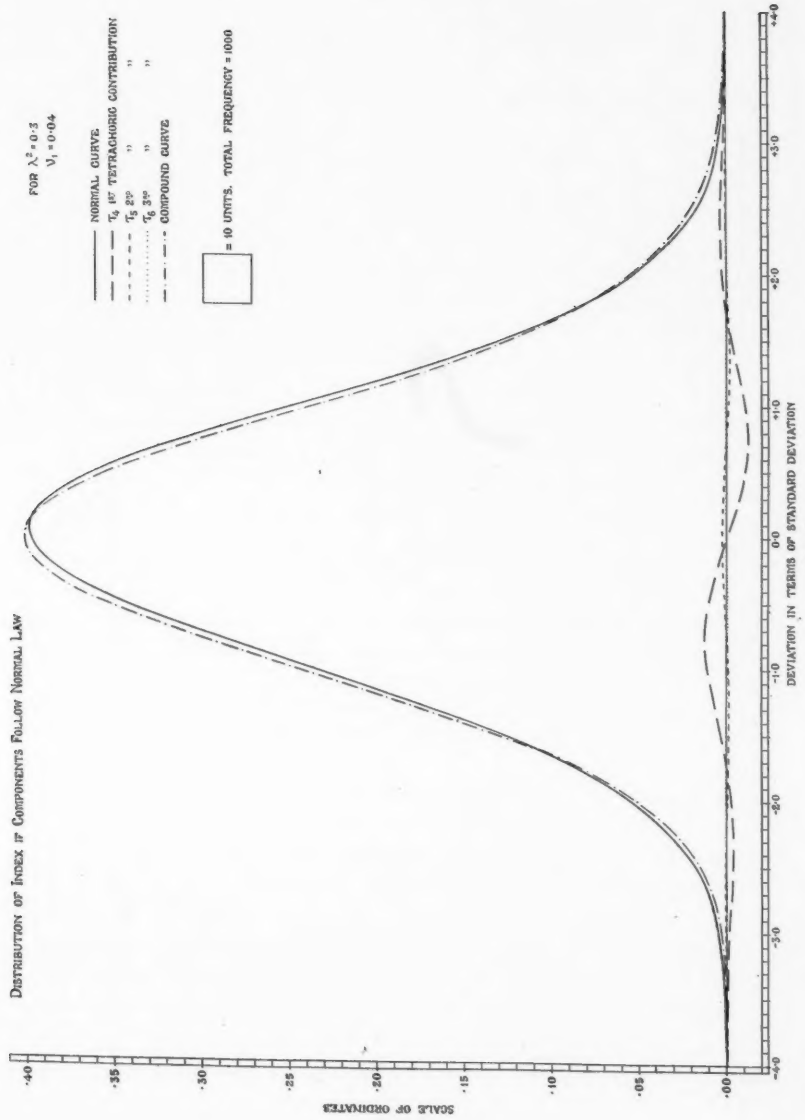

DISTRIBUTION OF INDEX IF COMPONENTS FOLLOW NORMAL LAW
FOR $\lambda^2 = 0{\cdot}3$
$\nu_1 = 0{\cdot}04$
NORMAL CURVE
γ_6 15° TETRACHORIC CONTRIBUTION
γ_6 22° " "
γ_6 30° " "
COMPOUND CURVE
= 10 UNITS. TOTAL FREQUENCY = 1000
SCALE OF ORDINATES
·40 ·35 ·30 ·25 ·20 ·15 ·10 ·05 ·00
-4·0 -3·0 -2·0 -1·0 0·0 +1·0 +2·0 +3·0 +4·0
DEVIATION IN TERMS OF STANDARD DEVIATION

Fig. 2.

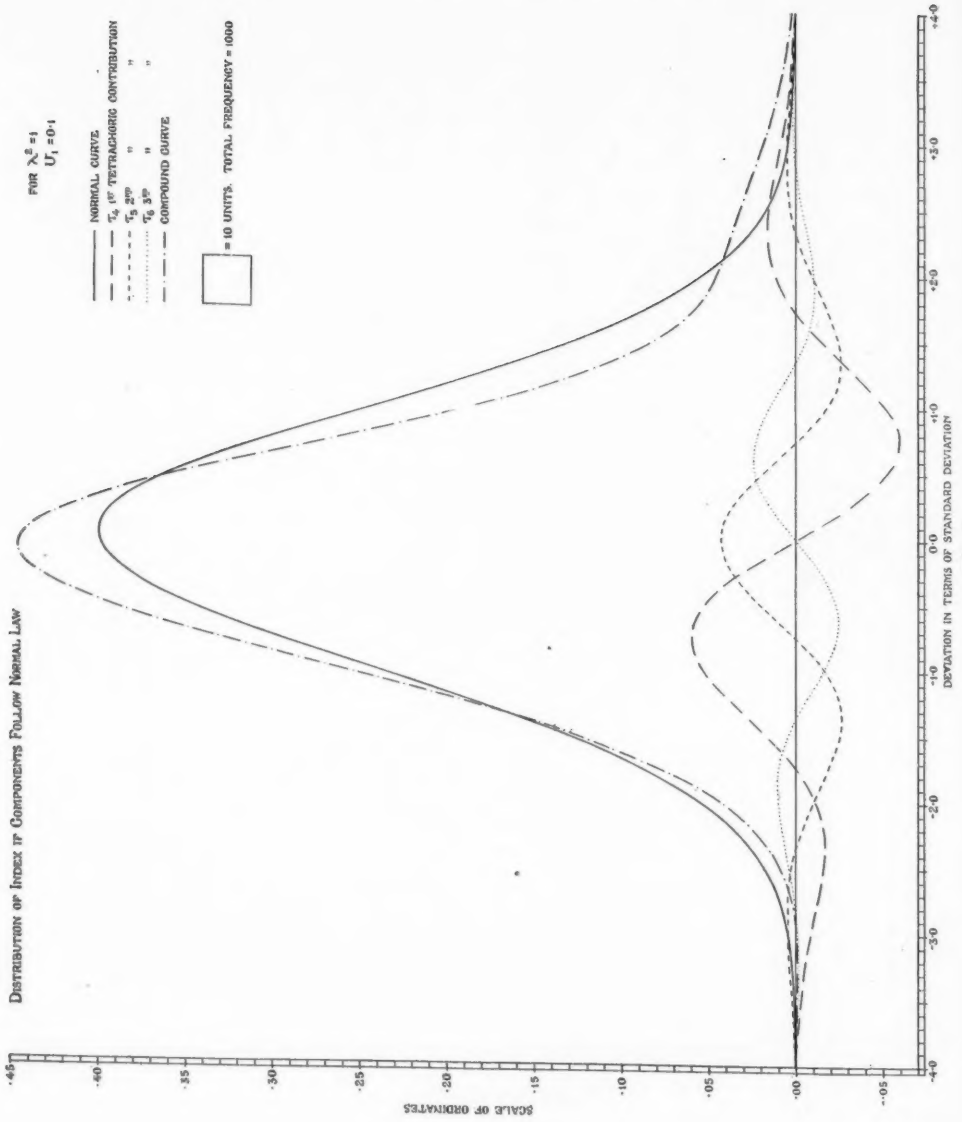

DISTRIBUTION OF INDEX IF COMPONENTS FOLLOW NORMAL LAW
FOR $\chi^2 = 1$
$U_1 = 0{\cdot}1$
NORMAL CURVE
τ_4 1ST TETRACHORIC CONTRIBUTION
τ_5 2ND " "
τ_6 3RD " "
COMPOUND CURVE
= 10 UNITS. TOTAL FREQUENCY = 1000
SCALE OF ORDINATES
·45 ·40 ·35 ·30 ·25 ·20 ·15 ·10 ·05 ·00 ·05
-4·0 -3·0 -2·0 -1·0 0·0 +1·0 +2·0 +3·0 +4·0
DEVIATION IN TERMS OF STANDARD DEVIATION

Fig. 3.

The sum of these and

$$y_0 = \tau_1 \quad \text{(the normal curve)}$$

gave the ordinates for the tetrachoric expansion (to the degree of approximation used).

In the figures are shown the curves of y_1, y_2 and y_3 (when possible) as well as y_0 and the composite curve.

Figure 1 shows the graph for $v_1 = \cdot 01$ and $\lambda = \cdot 316\,(\lambda^2 = \cdot 1)$. The τ_5 and τ_6 contributions were so small that they could not be shown upon the graph. The τ_4 contribution was also small, so the normal curve is a remarkably good fit for the curve of distribution.

Figure 2 shows the various curves for $v_1 = \cdot 04$, $\lambda = \cdot 548\,(\lambda^2 = \cdot 3)$. The τ_5 and τ_6 contributions are again small, but are shown. The curve of distribution still follows very closely the normal curve. Cases will seldom arise in practice in which the deviation from the normal curve will be much greater than that of Figure 2.

Figure 3, for $v_1 = \cdot 1$ and $\lambda = 1$, proves of little interest from a practical point of view, but illustrates well the unsatisfactory nature of the tetrachoric expansion*. Theoretically the case of $\lambda^2 = 1$ can arise only if $r = \pm 1$, a condition of affairs which is seldom approached in practice.

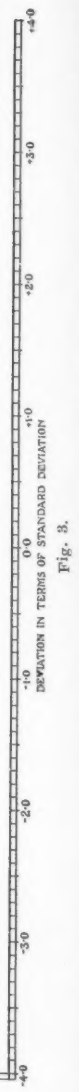


Fig. 3.

From the figure it is evident that the τ_4, τ_5 and τ_6 terms of the tetrachoric expansion all add considerably to the composite curve. The τ_5 contribution is somewhat smaller than the τ_4 contribution, but the contribution of τ_6 is not less than two-thirds of τ_5 at its maximum value. All three contributions are of approximately the same order. We can therefore only speculate as to whether the τ_7 and τ_8 terms might not bring further considerable changes to the composite curve. The convergence of the tetrachoric expansion is here too slow to justify the degree of approximation we have used, and there is no evidence that if the great labour of computing the coefficients of τ_7 and τ_8 were undertaken we should be any nearer the final goal.

It seems, then, that for conditions ordinarily met with in practice, the frequency distribution of the index, when both components follow closely the normal law, is sensibly normal. In other cases, however, particularly in cases where the correlation is high and the coefficients of variation are large, there may be a considerable deviation from normality.

[A further paper will shortly be published showing that comparatively slight deviations from the bi-variate normal distribution of x and y lead to marked divergence from normality in the index distribution, and that this is unfortunately the case in a number of anthropometric measurements. K.P.]

* Pointed out by Professor Pearson, *Biometrika*, Vol. XVII. (1925), p. 277.

ON THE EXAMINATION OF FINAL DIGITS BY EXPERIMENTS IN ARTIFICIAL SAMPLING.

By C. P. SUN.

In a recent paper* by Mr G. U. Yule some interesting data are given of the influence of personal equation in scale readings. That such personal peculiarities do occur, and sometimes occur with such intensity as to render useless long series of data, has been for many years a familiar fact to those whose work lies in anthropometry. On the other hand experience shows that inequality in the numbers of final digits may still exist even when the observer is highly trained and is reading to the nearest scale reading only, without estimating tenths of such a reading†.

In many measurements in anthropometry and biometry the readings can only be taken to a unit, say a millimetre, and the standard deviation may well be only a few millimetres. Grouping by millimetres will not give us more than 15 to 25 frequency classes, and it depends largely on where the scale of digits crosses these classes, which digits shall predominate. Mr Yule has referred to this matter in the paper cited above, p. 572, but as it is known to be a source of inequality in final digits, it is, perhaps, worth while noting what its effects may be experimentally. Let us take for example the upper face height on the skull measured in millimetres; the standard deviation is 4·mm. roughly. With a thousand individuals there would not be more than 20 to 25 groups; with a small cranial series probably only 15. With series of 15 to 17 crania such as those which were measured by Dr Morant in *Biometrika*, Vol. XIV., which Mr Yule cites, even many of these 15 to 20 groups will have no representatives at all and this produces further irregularity in the distribution of digits when they are obtained from measurements on short series.

Generally the mere fact of random sampling produces irregularities, often very considerable in the case of small samples, in the distribution of final digits, when the final digit is the unit of measurement and grouping.

That this inequality may be evened out if enough series of adequate length be taken—it does even out the more quickly, if our frequency distributions are symmetrical curves, and there are enough of them—I do not deny. And, as a

* "On Reading a Scale," *Journal of the R. Statistical Society*, Vol. xc. pp. 570—587.

† In dealing with Fawcett's and Macdonell's cranial measurements, it is difficult to see what Mr Yule has done, because those measurements were for lengths and circumferences taken to the nearest tenth of a millimetre by appreciation, and then this appreciation was used to class to the nearest millimetre, or in the case of some measurements to the nearest half-millimetre. Hence such deviations from equality of final digits are, we believe, due to the causes discussed in this paper.

matter of fact, I have treated this topic separately in two parts where I sought an explanation for the existing inequality and endeavoured to ascertain the extent to which symmetry and many groupings affect the frequency distributions of the final digits.

Part I.

If the series be relatively short, if some of the series be not symmetrical, but skew, if we allow for the usual sampling errors, will the result show even approximately an equality of final digits for such series and such units and such standard deviations, relative to the groupings by reading units, which we meet with in craniometry more especially, and not infrequently in anthropology?

This is the question upon which it seemed desirable to throw some light by aid of experimental sampling.

For this purpose Tippett's *Random Sampling Numbers** were used and five normal curve frequency distributions were obtained by the method indicated in the "Foreword" to that work. In precisely the same manner the frequency distributions for five skew curves were obtained by taking from the *Tables of the Incomplete* Γ-*Functions* as the sampled populations five distributions with different values of p. We have thus 10 distributions, with random variations from smooth continuous frequency distributions, partly symmetrical and partly skew, such as might easily arise in anthropometric practice. The normal curve frequencies were grouped in (a) 30 subranges, (b) 21 subranges, (c) 15 subranges, (d) 14 subranges, and (e) 10 subranges, and the corresponding numbers of individuals were (a) 300; (b), 152; (c), 300; (d), 152; (e), 300.

For the five skew curves the number of subranges were: (a'), 21; (b'), 16; (c'), 15; (d'), 12; (e'), 11 respectively; and the number of individuals in each frequency: (a'), 152; (b'), 200; (c'), 152; (d'), 200; and (e'), 152 respectively.

Such distributions might quite well represent millimetre measurements of ten characters on the skull, although the numbers are large for cranial series. Ten cards were now taken with the ten digits 0, 1, ... 9 inscribed on them, and drawn one at a time and replaced in the bag. Each draw settled the digit with which the frequency distribution was to start. The frequency of the final digits was then found to be given by Table I on p. 66.

Any one examining the totals column would assert that the observer had a marked prejudice in favour of the first five digits, 0—4. Yet the whole distribution is the result of random sampling from continuous symmetrical and asymmetrical distributions and of placing the scale down on the marginal frequencies by chance! If it be said that the number of subranges is too small, we may cite cases like the measurement of stature; no wise anthropometrician would record it to more than the nearest inch, because stature is capable of varying, possibly $\frac{1}{2}$ inch, according to the time of day at which it is measured. Thus, 1000 cases will only provide about 18 groups, and 200 possibly not more than 14 or 15. Forearm in adult women tabled to the *half*-inch, which is as close as can be reasonably

* *Tracts for Computers*, No. xv. Cambridge University Press.

ascertained, only gives 14 groups for 1000 cases, and for 200 cases might easily give only 10*. When we come to short cranial series, some of which Mr Yule makes

TABLE I.

Final Digit	(*a*)	(*b*)	(*c*)	(*d*)	(*e*)	(*a'*)	(*b'*)	(*c'*)	(*d'*)	(*e'*)	Totals
0	40	13	36	13	17	11	39	25	27	18	239
1	38	17	47	13	56	10	26	31	43	12	293
2	24	12	45	19	72	17	24	22	33	6	274
3	28	17	42	21	63	15	17	18	24	10	255
4	25	22	44	21	41	23	19	10	25	5	235
5	29	20	20	18	32	23	9	15	5	6	177
6	29	13	24	14	9	15	6	8	11	10	139
7	28	16	17	15	1	14	12	4	13	33	153
8	28	12	11	12	3	9	19	9	8	30	141
9	31	10	14	6	6	15	29	10	11	22	154
N	300	152	300	152	300	152	200	152	200	152	2060

use of with only 15 to 17 individuals, it is clear that we shall get down to 10 groupings, or the position of the scale against the distribution curve of the sampled population and the large variation of the random sampling for small numbers are factors of great importance in determining the distribution of final digits. Table II gives our results compared with the hypothesis of an even distribution of final digits. The coefficient of variation is 27·284 °/₀, and the $\chi^2 = 153{\cdot}354$; for such a value of χ^2 and for $n' = 10$ groups, it is generous to put

$$P = {\cdot}000{,}000;$$

in other words, the result is practically impossible on the basis of random sampling from equally probable final digits.

TABLE II.

Final Digit	Observed Frequencies f'	Theoretical Frequencies f	$f'-f$	$(f'-f)^2$	$(f'-f)^2/f$
0	239	206	33	1089	5·286
1	293	206	87	7569	36·742
2	274	206	68	4624	22·446
3	255	206	49	2401	11·655
4	235	206	29	841	4·082
5	177	206	−29	841	4·082
6	139	206	−67	4489	21·791
7	153	206	−53	2809	13·635
8	141	206	−65	4225	20·509
9	154	206	−52	2704	13·126
Totals	2060	2060	—	$\sigma = 56{\cdot}21$	$\chi^2 = 153{\cdot}354$

* The temperature of normal individuals can hardly be read to advantage closer than 0·1°, and 100 normal cases will only give 10 groups.

In the above investigation we have taken rather long series for craniometric measurements, 150 to 300, thus reducing the variations due to random sampling, but this example seems to suggest that one must be very cautious in assuming that every case in which the final digits of an observer are very unequal and give an impossible "P" is due to personal predilections. The illustration may also account for deficiencies or excesses of certain digits which can hardly be the subject of prejudice at all.

Part II.

The first part of this investigation has demonstrated the distribution of final digits as they may occur in craniometrical and anthropological works with relatively short series and when each group corresponds to a digit; and the present part attempts to show what changes it would make, if we confined our attention entirely to those distributions which have a fairly large number of frequency classes. In precisely the same manner as before, three normal curves and three skew curves not identical with the previous ones given in Part I were obtained. The number of subranges for the normal curves is: (a), 23; (b), 28; (c), 27; for the skew curves: (a'), 26; (b'), 24 and (c'), 25 respectively.

Such distributions might represent anthropological units. Yet the distribution of the final digits resulting from mere chance, and without personal predilection, is shown by the following table:

TABLE III.

Final Digit	(a)	(b)	(c)	(a')	(b')	(c')	Totals
0	23	20	10	10	23	6	92
1	17	24	12	22	19	9	103
2	14	23	10	16	16	24	103
3	9	16	14	19	21	13	92
4	15	16	9	12	12	14	78
5	16	14	16	23	15	15	99
6	13	13	16	20	20	11	93
7	18	26	11	10	35	13	113
8	18	28	8	10	24	15	103
9	7	20	19	8	15	5	74
N	150	200	125	150	200	125	950

Examining superficially the data, one would say that in a distribution like this, there was a special predilection for 1, 2, 5, 7, 8; indifference toward 0, 3, 6; and unusual prejudice against the digits 4 and 9. This we hold has no basis at all, but is a matter of pure accident. Here again we meet with inequality of final digits, even with more groupings and far greater frequencies than are usual in craniometrical measurements.

We may again compare this distribution by aid of the (P, χ^2) test with the even occurrence hypothesis. The value of χ^2 is 13·4943, which gives a P value of ·1427.

Relatively speaking, it is more probable than the results recorded in Part I, but it is by no means markedly significant. Since we suspect skewness must have played an important rôle in changing the predominancy of certain digits, the χ^2's for the symmetrical and the asymmetrical distributions were calculated independently; and they are 6·788 and 16·979 for the normal and skew curves respectively. These give the corresponding P values of ·6589 and ·0491, the probability of occurrence being quite good for the normal curves, thus indicating concisely what symmetry can do in increasing the chances of even occurrences.

As the starting of the digits is perfectly arbitrary, we might have summed up all the observed frequencies both in Part I and Part II, maintaining the same serial order, thus:

276	232	266	244	228	331	396	377	347	313
4	*5*	*6*	*7*	*8*	*9*	*0*	*1*	*2*	*3*

when it would have been explained as a case of prejudice in favour of the whole number, or its immediate neighbourhood, if our last digit was a decimal figure. Accordingly, while fully realising in how many cases careless or prejudiced readings are an important factor in the inequality of the numbers of final digits, we believe that other factors come into play which even the most careful and physically trained measurers cannot avoid, and that Mr Yule, while recognising their existence, has possibly not sufficiently emphasised them. This emphasis is provided by the above experimental sampling, which is not far removed from cases actually occurring in anthropometry, and in particular craniometry, although in the latter case the total frequencies must be considered as exaggerated.

SIMILARITIES AND DIFFERENCES IN THE SPECIES OF THE MULBERRY SILK-WORM*.

By Professor A. V. ANOOCHIN, Moscow.

Silk-worm breeders and dealers in raw silk have for long been greatly interested in the possibility of a comparison of the different species of the mulberry silk-worm, the aim being to find out which species is the most profitable and to what relative extent.

Silk-worm breeders approach this question quite simply. The species are compared on the basis of a large amount of data and then conclusions are drawn.

Here is an example of such an investigation. Ivanoff†, comparing the qualities of the Baghdad species with others, writes much as follows:

In its percentage contents of silk in cocoons the Baghdad variety stands slightly below the European species, occupying the fourth place, as seen from the following tables which represent the standard numbers obtained in the last ten years in the laboratory of the Caucasian Silk-worm Breeding Research Station.

TABLE I.

Percentage Contents of Silk in Cocoons.

		Silk
1	Chinese White	36—35 %
2	European „	35—34
3	Japanese	34—33
4	Baghdad	32·5
5	Middle-Asiatic	32
6	Persian	30
7	Caucasian	29

In all other respects it is either superior to European species or almost equals them, which is shown in the following tables:

TABLE II *a*.

Length of Thread in a Cocoon (the averages of ten years' research).

		metres
1	Baghdad	866
2	European	724
3	Middle-Asiatic	723
4	Persian	717
5	Caucasian	639
6	Japanese	580
7	Chinese	546

TABLE II *b*.

Strength of Thread (expressed by weight in grams of Silk-thread of a given Length).

		grams
1	Caucasian	38·0
2	Middle-Asiatic	36·0
3	Chinese	35·8
4	Japanese	35·1
5	Persian	33·6
6	Baghdad	33·4
7	European	33·3

* Memoir prepared in the Central Silk Experimental Station, Moscow; Director, B. Michin.

† V. P. Ivanoff, "On the question of the Baghdad Species and its Replacement by European and Crossed Species." *News of the Experimental Silk Station*, 1915, Tiflis.

TABLE II*c*.

Elasticity.

(Extensibility °/₀)

1	European	21·0
2	Middle-Asiatic	20·8
3	Baghdad	20·4
4	Persian	20·2
5	Caucasian	20·0
6	Japanese	19·0
7	Chinese	18·6

TABLE II*d*.

Thickness of Thread.

1	Chinese	137
2	Caucasian	139
3	Japanese	149
4	Middle-Asiatic	153
5	Persian	159
6	European	172
7	Baghdad	173

Thus far Ivanoff.

Amongst the Moscow manufacturers one hears the opinion that the thickness of the silk-thread of the Baghdad variety is very uneven, so that it is impossible to spin it into an even thread. Laboratory investigations do not confirm this. To prove this statement I mention the results of examination of three species: Ascoli Yellow from the firm Paskvaliss, Crossed Spherical—firm Paskvaliss and Baghdad White—firm Pinachi. The cocoons were obtained from the Caucasian breeders and were of their own breeding and from the same district.

For each variety five average cocoons were taken; the thickness of the thread is determined by the weight of each hundred metres expressed in milligrams.

TABLE III.

Weight of each hundred Metres expressed in Milligrams.

a. Askoli Yellow. *b*. Crossed Spherical. *c*. Baghdad White.

Length in Hundreds of Metres.	*a* Cocoon					*b* Cocoon					*c* Cocoon				
	A	*B*	*C*	*D*	*E*	*A*	*B*	*C*	*D*	*E*	*A*	*B*	*C*	*D*	*E*
1	30	35	35	35	25	45	40	30	40	25	45	35	40	40	35
2	35	40	35	40	30	45	35	35	40	35	40	35	40	35	35
3	35	40	35	35	30	45	30	30	30	35	40	30	35	35	35
4	30	35	30	35	30	40	25	25	25	30	40	30	35	30	35
5	20	25	25	25	30	40		25	15	30	35	25	30	30	35
6		20	25	20	30	40				25	35	25	25	30	30
7			20		30	30				25		20	20	25	30
8					25					20			20	20	25
9					25										

From these facts we can by no means conclude that the variation in weight of the separate silk-threads of the Baghdad variety is greater than that of the first two varieties, which are considered to be the best Italian.

What is the general deduction from this? From all these facts one concludes that in its technical qualities the silk of the Baghdad species stands by no means so low as some Moscow manufacturers maintain.

The conclusion is quite unexpected. In fact in length of thread it holds the first place amongst other species, in elasticity the third, in percentage contents of silk the fourth, in strength of thread the sixth, in thickness the seventh. Are then

its qualities superior or inferior to those of other species? One does not know. In one quality it is better, in another worse, and it is impossible to draw a final conclusion for two reasons:

The first lies in the fact that one cannot compare species according to separate characters. One always obtains the result which is so well expressed in the following words of Serebrovski:

"On account of transgressive variation it is impossible to select a character which would serve as an absolute criterion for the distinction of one species from another. According to one of its characters the individual resembles more closely one definite species, according to another it may resemble more closely a second species*."

The second reason we have already discussed in our work, "The Changeability of Silk-Thread†," wherein it was said that in order to find the standard characters which define a species it is necessary to examine the greatest possible number of individuals. With a few individuals we may commit great mistakes, since every character forms a variational series and accidentally selected individuals may either correspond to the standard type, or may diverge from it on the positive or negative side. Therefore such individuals should not be compared with each other, especially in the case where they come from different species. Our results‡, from the comparison of the white Baghdad variety with the Askoli and Chinese-golden, bring us to quite a different conclusion, since we find that, according to all characters, except thinness of thread, the Chinese variety occupies the third place.

TABLE IV.

	Frison*	Weight of Thread	Length of Thread	Thinness of Thread	Bassinet†
Baghdad	41·93	253·43	757·58	·34	31·79
Askoli	35·84	205·52	781·27	·27	16·40
Chinese-golden ...	19·55	187·66	607·26	·31	10·81

* The flock silk or curly which must be taken off before the end of the silk filament for reeling can be found.

† When the reeling is finished the pupa remains enveloped in a silk enclosure, the "husk," which is not reeled off and is called the bassinet. The bassinet varies with the breed of silk-worm in weight.

In spite of these apparent indications of the possibility of comparison between the three mentioned species, we still cannot draw any conclusions, since on the one hand the Baghdad variety occupies the second place in length of thread, while on the other, if we increase the number of species, we shall become confused in the estimation of their qualities.

* Serebrovski, "The Statistical Method in Biology," in "The Statistical Method in scientific investigation." *Academy*, 1926.

† A. V. Anoochin. Central Silk Research Station, 1927.

‡ A. V. Anoochin, "The Changeability of Silk-Thread." Central Silk Research Station, 1927.

The only proper method is the one based upon the laws of large numbers. In which case also the degree of similarity ought to be expressed by a single coefficient embracing all the most important characters of the two species under examination. Such a coefficient would furnish the means of finding the degree of likeness of two or more species, which approach each other nearly or remotely, and of calculating by how much the degree of likeness of any two exceeds the degrees of the others.

The first attempt of this kind gave brilliant results, and was that of Heincke. In his monograph, *The Natural History of the Herring*, Heincke put before himself a whole series of questions, one of which was the following: Does the herring of the northern European seas represent one species or is it a group of local species which can be clearly distinguished from each other? A question extremely difficult to solve since these herrings possess very similar features, and although they are subdivided into species these latter approach each other very closely. Heincke applied the method of variational statistics and set himself to study the variations in separate characters in the individual herrings; he performed 100,000 measurements and found from these that the characters of the individual herring vary according to the law of error and give a typical curve. Of such characters there were 60. If one compares the curves of the characters of individual herrings with those of the supposed species it becomes apparent that each species has its own average value for every character. It follows that for the definition of a species it is sufficient to take the combination of all standard (average) values of the different characters. When this was done one obtained the expected result, i.e. the greater the geographical distance between the species the greater are the differences between them.

The next task was still more difficult—that was to find a means of distinguishing one species from another. By aid of only one character it is impossible to achieve this, as explained above. It is necessary to bear in mind the combination of all characters in order to do it. Heincke's method was the following. He found that every variational series is characterised by the fact that the sum of the squares of deviations from the average value is minimal, i.e. if we compare the sum of the squares of deviations from the average value with the sum of the squares of some other number, the former will always be the smaller. From this it follows that the animal in question will belong to that species, with respect to whose average values the sum of the squares of the deviations of all its characters will be minimal.

Thus, if we calculate the sum of the squares of the deviations of a certain individual from the average values of the same characters of all known herring species, the smallest number will indicate to what species the individual belongs.

If, therefore, at the end of the calculation the following results be obtained:

for the White Sea Herring.	3·213
„ „ Spring Norwegian Herring . . .	3·696
„ „ Jutland Herring	6·317

it means that the individual in question belongs to the White Sea variety.

This method of Heincke has not met with wide application, apparently for the reason that it involves an enormous amount of measurement and calculation. We also hesitated to follow it and waited for help from the biometric-mathematicians to simplify it. At the same time we regard the method devised by Heincke as the most correct*; the problem is to diminish the amount of measurement and calculation.

In the search for such a coefficient we met with the work of Pearson, "On the Coefficient of Racial Likeness," *Biometrika*, Vol. XVIII. Parts I and II, 1926, p. 105. This coefficient has been proposed by him for craniologists and just answers our requirements.

"Every craniologist and indeed every physical anthropologist has come up against the difficulty of comparing two races of which it is only possible to secure a limited number of individuals of one or other or both races. Not unnaturally he is driven under the circumstances to seek help by measuring a large number of characters in order to compensate for few individuals. We have frequently to admit that relatively few individuals are available in many anthropometric inquiries, and that we really must compensate for the smallness of our sample by the largeness of our character series."

"But how is this to be done? We can compare the means for our two small groups character by character, and if we are trained statisticians we shall compare these mean differences with their standard deviations. But when a considerable number of the characters do not show differences markedly significant with regard to their probable errors, we are left in considerable doubt as to what inference may be safely drawn from the whole series. We need a single numerical measure of the whole system of differences, something which will express by a single coefficient the measure of resemblance (or divergence) of the two races or groups. Such a measure or coefficient I term a *Coefficient of Racial Likeness* (C.R.L.). It should be a measure, not of how far the two races or tribes are alike or divergent, but of *how far on the given data* we can assert significant resemblance or divergence."

Pearson's Coefficient of Racial Likeness has the following numerical expression:

$$\text{C.R.L.} = S\left\{\frac{1}{M}\left(\frac{(m_s - m_{s'})^2}{\frac{\sigma_s^2}{n} + \frac{\sigma_{s'}^2}{n'}}\right)\right\} - 1 + \left[\frac{1}{M}\right] \pm \cdot 67449\sqrt{\frac{2}{M}}.$$

m_s and $m_{s'}$ are the averages of the same characters of two species, σ_s and $\sigma_{s'}$ their standard deviations, n and n' the number of individuals, M the number of characters and S the sum of values in the large brackets for the whole number of characters. The ideal value of M is, according to Pearson, 40—50, but in practice 24—25 is quite sufficient for small values of n, n'. In any case the greater the values for n, n', the smaller the error, and the more correct the coefficient.

* [I venture to think the method essentially erroneous. It could not be applied to characters measured in different units and it entirely neglects the differing variabilities of the several characters, which are therefore improperly weighted, the larger characters being weighted with their size. K. P.]

We calculated the C. R. L. for cocoons of the Baghdad white, Askoli and Chinese-golden variety on the basis of eight characters (longitudinal diameter, transverse diameter, weight of complete cocoon, weight of the cocoon-envelope, weight of frison, weight of bassinet, weight of silk-thread and length of silk-thread). The small number of these was compensated by the great number of individuals, 250 cocoons.

As a result we obtain the following numbers:

TABLE V.

Coefficient of Racial Likeness.

	Baghdad	Askoli	Chinese-golden
Baghdad	—	396·151 ± ·337	565·230 ± ·337
Askoli	396·151 ± ·337	9·792 ± ·380*	283·256 ± ·337
Chinese-golden ...	565·230 ± ·337	283·256 ± ·337	—

* One group of Askoli from Astrachan, the other from Moscow: see below, p. 75.

Great numerical values of the coefficient made us mistrust its practical use, the more so, since when applied by craniologists its numerical expression never grows beyond two digits, even in the case when the difference between the characters of the compared skulls is very great. Besides this, large numerical expressions are inconvenient†.

For the sake of verification we introduced a ninth character, the thinness of thread. As a result of this enlargement the numerical value of the coefficient at once diminished considerably.

TABLE VI.

	Baghdad	Askoli	Chinese-golden
Baghdad	—	374·023 ± ·318	512·433 ± ·318
Askoli	374·023 ± ·318	—	257·709 ± ·318
Chinese-golden ...	512·433 ± ·318	257·709 ± ·318	—

The decrease of coefficient indicates that with a greater number of characters its numerical expression will become smaller and more correct, since parallel with this the amount of error decreases, depending entirely on the number of characters considered.

† [Similar large values have been obtained when the coefficient is applied to measurements on the living of different races. Even allowing for the great differences of technique in such measurements, the high values seem to indicate that racial differences lie to a greater extent in the fleshy parts than in the skeletal frame of Man. K. P.]

Pearson considers 40—50 examinations as ideal and 24—25 as sufficient. But we have only considered nine characters, therefore we obtained a greater numerical value for the coefficient. Pearson points out that in the case where one selects "the most diverging or the nearest alike characters" for two species, the amount of these does not affect the C.R.L.

In order to verify that the coefficient really does show up the likeness or divergence of species, we calculated it for two groups of cocoons of the Askoli variety (each group 250 cocoons). One group of cocoons was obtained from Astrachan dating from 1922, the other dating from 1926 was received from the Central Silk-worm Breeding Station (Moscow). The following six characters were selected for comparison: the longitudinal and transverse diameters, the diameters of the two bulgings on the cocoon, the complete weight of a cocoon and the weight of the envelope. The coefficient was equal to $9{\cdot}792 \pm {\cdot}380$.

According to Pearson the coefficient fluctuates round zero with the probable error of $\cdot 67449\sqrt{\frac{2}{M}}$, when both groups have been taken from the same species. Our C.R.L. is greater than zero, apparently because only a few characters were considered. But it differs so definitely from the C.R.L. of two different species, that one can recognise its use for judging likenesses and differences between the species of the mulberry silk-worm as a proved matter*.

In his work, Pearson points out that "the greater the number of skulls the greater the coefficient," and before that, "I next proceeded to consider what influence the number of individual crania dealt with had on the coefficient. Considering only coefficients based on all characters (i.e. absolute size and indices and angles), I obtained a correlation table of 580 entries, entering each coefficient twice, once for each race dealt with. The table was still more 'lumpy' than the previous one, as the skull frequencies run from 6 to 885. The answer I found for the Correlation of the C.R.L. and Number of Crania used was

$$r = + {\cdot}1270 \pm {\cdot}0276,$$

i.e. there was a not very important correlation between the number of skulls used and the coefficient, the *greater* the number of skulls the *larger* the coefficient†."

We have obtained such high values for our coefficients, since we chose a fairly large number of individuals, 250 cocoons. The diminution in the number of individuals has for us a threefold meaning:

(1) To obtain a smaller coefficient in the numerical sense.

(2) A considerable diminution of measurement and calculation.

(3) If the obtained coefficients are in the same ratio to each other as when calculated from a small number of cocoons, we shall have one more proof of the

* [Even the coefficient 9·792 is probably explicable on the basis of differentiated Askoli races, either by inheritance or long continued environment. K. P.]

† *loc. cit.* p. 116.

correctness and possibility of practical application of the C.R.L. to the species of the mulberry silk-worm and therefore to species of any other animal.

For this purpose we undertook a final verification of the formula and calculated the C.R.L. for the same species and for the same nine characters, but for any casually chosen ten cocoons, and obtained the following results:

TABLE VII.

	Baghdad	Askoli	Chinese-golden
Baghdad	—	50·313 ± ·318	110·295 ± ·318
Askoli	50·313 ± ·318	·505 ± ·380	37·731 ± ·318
Chinese-golden ...	110·295 ± ·318	37·731 ± ·318	—

The coincidence is striking. The coefficients preserved their ratios and moreover the C.R.L. for other groups of Askoli has received the numerical value which Pearson attaches to it—it is near to zero. We think that the latter fact verifies most conclusively that the formula for Pearson's C.R.L., which is employed by craniologists, can also be used for the purpose of defining the likenesses and divergences of species.

As a result of the analysis by aid of the formula for the C.R.L. we can draw the following conclusions:

(1) The likeness is greatest between the Askoli and Chinese varieties; the least between the Baghdad and Chinese varieties; the Baghdad variety is nearer to the Askoli than to the Chinese-golden variety.

(2) Taking arbitrarily the Chinese-golden variety as the standard one, one can say according to (1) that the Askoli occupies the second place after the Chinese species and the Baghdad species the third, i.e. the latter is the worst of the three species examined.

(3) According to the preceding statements (1) and (2), or from the comparison between the Baghdad and Askoli varieties with the Chinese (the first and second ratios), it follows that the Baghdad variety diverges from the Chinese three times further (110:38) than the Askoli from the Chinese, which can be seen from the following numbers:

			C.R.L.	Degree of Likeness
1	A. . .	Ch. . .	38 . .	1
2	B. . .	Ch. . .	110 . .	3·0

(4) If one calculates the C.R.L.'s for all species of the mulberry silk-worm and places these in a certain order, one obtains a genealogical table which may solve indirectly the problem of the origin of species and show the degree of their relationships. For this purpose one may compose, as is done by craniologists, a table of relative degrees like the one proposed by Pearson for human races*.

* *Biometrika*, Vol. XVIII. p. 112

(5) Such a table of degrees of likeness furnishes a means of orientation for the origin of species from each other and provides aid in investigating it. If for instance the C.R.L. of two species is very small and these species have only one strongly divergent character (for instance, thinness of thread), the investigation of their origin ought to be carried out in one of the following three ways: Find whether they have one common ancestral form with a lesser thickness of thread, or whether there are intermediate varieties giving a gradual transition from one to the other which we do not yet know, or finally whether one sprang from the other by the process of mutation.

For the solution of these problems it is necessary, as already mentioned, to find the C.R.L. for as many species as possible. This involves a large amount of measurement and calculation, and it is necessary to standardise the list of characters, their quality (which particular character) and the number of cocoons.

(6) Practice suggests that cocoons of the first generation of crossed parents are better than the parent cocoons. The determination of the C.R.L. of these cocoons and of either parent species will show us the likenesses and differences between these cocoons and also to which parent species they are more closely related. If the difference is great, one ought to analyse in detail the cocoons of the first generation, if it is not there is no need for this, since the cross-bred cocoons differ only slightly.

(7) The C.R.L. gives us the means of finding the origin of an unknown species. If we discover cocoons of such a species and suppose that according to its appearance it resembles some Eastern species, for instance the Baghdad species or Buchora species, we only have to calculate the C.R.L. for the unknown species and for the supposed related species in order to find which species is most nearly related to it. This, in its turn, will enable us to judge its qualities.

(8) Having arbitrarily chosen some species as the standard one and compared with it any other species, the quality of which we want to determine, we make use of the C.R.L., determining it for each of the species, and thus find the likeness and the divergence. After having thus compared a whole series of species with the standard one, using a standard number and quality of characters and a standard number of cocoons, we can compose a table of C.R.L.'s placed in ascending order of magnitude. Such a table makes it possible to find how great a likeness or divergence of certain species may be, as already mentioned under paragraph (3) of this Summary. This possibility is the result of the fact that the numerical expression, the C.R.L., loses its abstract character and obtains a concrete meaning, showing the degrees of similarity and difference.

In conclusion, I may put on record the means (M), the standard deviation (σ), and the total frequencies used in computing the above coefficients of racial likeness, which may be found useful for further investigations.

TABLE VIII.

	Unit of Measurement	Baghdad			Askoli			Chinese-golden		
		M	σ	*n*	*M*	σ	*n*	*M*	σ	*n*
Longitudinal diameter ...	mm.	39·146	2·546	250	32·180	2·561	250	26·126	2·239	250
Transverse diameter ...	,,	18·114	1·628	,,	15·410	2·247	,,	17·407	1·493	,,
Weight of complete cocoon	mg.	1·655	·430	,,	·624	·106	,,	·487	·160	,,
Weight of cocoon-envelope	,,	·322	·620	,,	·264	·045	,,	·175	·039	,,
Weight of frison ...	,,	41·932	21·511	,,	35·844	20·885	,,	19·550	11·853	,,
Weight of thread ...	,,	253·427	50·294	,,	205·520	46·900	,,	187·655	36·935	,,
Weight of bassinet ...	,,	31·785	13·923	,,	16·398	8·826	,,	10·810	2·802	,,
Length of thread ...	mm.	757·576	158·120	,,	781·272	159·210	,,	607·261	125·197	,,
Thinness of thread ...	mm./mg.	·340	·060	,,	·270	·053	,,	·310	·055	,,

DEXTRALITY AND SINISTRALITY OF HAND AND EYE.

SECOND MEMOIR.

By T. L. WOO, B.A., National South Eastern University, Nanking,
M.A., Columbia University, N.Y.

CONTENTS.

	PAGE
I. Introductory	79
II. Source of Data and Methods of Treatment	80
III. Problems and Results:	83
1. For every character of hand and eye as a whole, which side is dominant?	83
2. Is Manual Laterality absolute?	85
3. Is Ocular Laterality absolute?	85
(*a*) On the Relationships of Percentage Occurrences of the Dominance of Ocular and Manual Characters	85
(*b*) On the Quantitative Interrelations of the Dominance of the Ocular Characters	101
(*c*) On the Quantitative Interrelations of the Dominance of the Manual Characters	108
4. Is there any Truth in the Theory of Universal Laterality?	121
(*a*) On the Relationships in Percentage Occurrences between the Dominance of Ocular and Manual Characters	121
(*b*) On the Relationships in Intensity between the Dominance of Ocular and Manual Characters	127
5. Is Laterality a Continuous Variate?	140
6. To what extent is the Correlation of Lateralities a Function of the absolute organic correlations?	143
IV. Summary and Conclusions	146

I. Introductory.

In a preliminary study of this subject* based on data collected by Sir Francis Galton, it was shown that the theory of "absolute laterality," or the theory that sinistrality in one character means sinistrality in all, certainly did not apply to left-handedness as measured by grip and left-eyedness as measured by acuity of vision. Not only were the percentages extremely different for sinistrality of hand and eye, but there was a zero correlation between the difference of grip of the right and left hands and the difference of acuities of the right and left eyes.

* *Biometrika*, Vol. xix. pp. 165—199.

The topic was there discussed on the basis of a very large number of measured individuals, 5000 to 7000, but only on the basis of the two above-mentioned characters.

The present paper carries the matter further by dealing with a far larger number of both manual and ocular characters, but on a far more limited range of subjects, about 400.

In this paper it seems unnecessary to repeat the history and criticism of the various theories of the subject, which appeared in the former memoir. We shall confine ourselves to an endeavour to answer the following questions.

(1) For the population as a whole, for each character of hand and eye, which side is dominant and by how much?

(2) Is the hand dominant for one manual character also dominant for a second? Is manual laterality absolute?

(3) Is the eye dominant for one ocular character also dominant for a second? Is ocular laterality absolute?

(4) Does the side which provides the dominant eye for one or all ocular characters provide the dominant hand for one or all manual characters? Is the theory of absolute laterality true for any or all characters?

(5) Is either manual or ocular dominance in itself a clear-cut characteristic, i.e. can we always say that a particular individual is either dextral or sinistral for a given character? Is ambilaterality non-existent or even of very rare occurrence, or does it occupy its appropriate place between the extreme values of dextrality and sinistrality? That is to say: Is laterality a continuous variate?

(6) To what extent is the degree of laterality, i.e. the difference of the character on right and left sides, a function of the intensity of the character itself?

II. Source of Data and Methods of Treatment.

In order to throw light on the above questions, the data for six manual and five ocular characters which are measured on students tested in the Anthropometric Laboratory of the Department of Applied Statistics, University College, London for both right and left sides, have been placed at my disposal by the Director of the Laboratory. The numbers of individuals available, all males, vary round 400 according to the character; a certain number of students not having presented themselves for all the tests. The ages run from about 18 to 24, but we have not thought it needful to make any correction for age for this range of years. It has been shown that the changes in ocular and manual characters for these ages are relatively small. The characters dealt with will now be specified:

1. Ocular Characters.

(1) *Near Point.* This test is measured in inches by the nearest distance at which a diamond type can be read clearly. The type is advanced towards the eye by steps of $\frac{1}{2}$ inch and the last $\frac{1}{2}$ inch interval from the eye at which the type can

be read is recorded. Clearly, the smaller the distance recorded, the better the subject's eye can accommodate itself. If the right side record in Near Point is larger than the left, it means the left eye dominates the right one in this respect.

(2) *Visual Acuity.* Instead of the Diamond Type used by Galton for Visual Acuity, the University College Anthropometric Laboratory uses Snellen's test types with additional lines of type to make a finer gradation. The subject sits at the distance of six metres. The score is taken from the last line read completely plus number of letters read after it. The number of letters in each line being known the score is calculated to the nearest second place of decimals as follows:

Scores	$\frac{6}{60}=\cdot10$,	$\frac{6}{36}=\cdot17$,	$\frac{6}{24}=\cdot25$,	$\frac{6}{18}=\cdot33$,	$\frac{6}{12}=\cdot50$,	$\frac{6}{9}=\cdot67$ and so on
Number of letters in line	1	2	3	4	5	6

For instance, if line $\frac{6}{24}$ is read completely and plus five letters *anywhere* beyond it, the score is taken $=\cdot25+\frac{4}{4}(\cdot33-\cdot25)+\frac{1}{5}(\cdot50-\cdot33)=\cdot36$. Our measures are obtained from the differences of distant vision between the two eyes.

(3) *Spherical Refractions.* (Refractions in the "horizontal"* meridian and in the "vertical" meridian.) Refractions are tested by the ophthalmoscope at one metre distance from the eye. No cycloplegic is used, but the subject is asked to gaze at a dimly illuminated card at six metres distance. Lenses are introduced until there is no movement of the "shadow." Two measures are taken for each eye for refractions in the "horizontal" and in the "vertical" meridian. To avoid the difficulty of scoring we have omitted those cases, very few in number, where right and left eye have refraction of opposite sign in the same meridian. The better eye is to be taken as the one having refraction least in amount in dioptres regardless of sign, the other eye is taken to be inferior by the difference in dioptres between the two amounts of refraction in a given meridian.

(4) *General Astigmatism.* Amount of astigmatism is measured by the difference in refraction to the nearest $\frac{1}{4}$ dioptre between axes of maximum or minimum refraction, i.e. difference between refractions in the two meridians classed as horizontal and vertical. The better eye is to be taken as the one having astigmatism least in amount, regardless of sign, and the other eye is inferior by the difference between the two amounts of astigmatism.

2. Manual Characters.

(1) *Grip.* Strength of grip is measured in pounds for each hand by a grip dynamometer. Three trials with each hand are allowed, introducing a brief pause between each trial to avoid excessive fatigue. For our purpose only the highest record registered at each trial for each hand was used.

(2) *First Pull and Last Pull.* Measured also in pounds by means of a dynamometer for the right and the left hands. The experimenter adjusts the instrument to height of subject's hand and instructs him that he is, at the signal

* Principal axis nearer the "horizontal."

"now," to pull as forcibly as possible, and to maintain this pull with his utmost effort until told to stop at the end of 30 seconds. The experimenter, after saying "now" for subject to begin, immediately takes the first reading, and thereafter glances at the scale and records every 5 seconds; thus seven readings for each hand are taken altogether. For our present purposes only the first and last readings, described briefly as measures of first pull and last pull, are chosen.

(3) *Difference of Strength at First Pull and Last Pull, or Lack of Endurance in Pulling.* It will be instructive to compare the initial strength with the final for each hand to obtain an index of persistency. For this purpose, we subtracted the last reading from the first and used the difference, which was taken in pounds, as a measure of the Lack of Persistency or Endurance.

(4) *Steadiness of Hand Test.* The subject stands by a table on which are placed two parallel horizontal brass plates at a distance of about 4 inches one above the other. The upper plate is pierced with holes, twenty in number, and lettered *A* to *T*. The subject is given a steel needle 1 mm. diameter wired to a battery, and holding it truly vertical, he has to pass it through the holes in the upper brass plate, touch the lower plate, when an electric bell rings, and extract it again without touching the wall of any hole; if he does touch, another bell of different tone is rung and the subject loses that hole. He starts with the hole *A* of $\frac{32}{64}$th inch diameter, and if he succeeds, he tries hole *B* of $\frac{30}{64}$th inch diameter, and so on. The last hole the subject can pass the needle through ringing only the first and not the second bell is a measure of his steadiness of hand. Each subject is allowed two trials. The marking starts at one for *A* and increases by unity up to the hole *M* at 13 marks, and there for the very fine holes increasing by two marks from *N* at 15 to *T* at 27. The marking depends solely on the last hole reached without contact with upper plate.

(5) *Balancing Test.* A flat piece of board is placed with one end slightly projecting over the edge of the table. Upon this board one or other of two vertical rods attached to discs of different small diameters can be placed. The subject has to take the board with right or left hand and raise it shoulder high with the arm outstretched and return it in the same way to the table without upsetting the rod. He makes three trials with each rod and with each hand, or 12 tests in all. He gets three marks for each success with the rod on the larger base and five marks for each success with the rod on the smaller base. The subject can thus obtain 0 to 24 marks with each hand. To obtain a continuous distribution, subranges of three were taken.

It should be noted that the Steadiness of Hand and Balancing Tests are not purely manual tests for they include some ocular control; they are really mixed tests, included under manual tests merely for convenience.

For all the above characters the intensity of laterality was measured by the difference of the performance of the right side from the left (right minus left); if this be positive the subject is termed a dextralist for this character; if negative

a sinistralist. But in such cases as Near Point, the two Refractions and General Astigmatism, the signs of individual scores are reversed (or left minus right) because we consider that the better score should be less in amount.

For simplicity we shall generally use the symbols, $A_1, A_2, \ldots A_6$ to represent the differences of the six manual characters, $B_1, B_2, \ldots B_5$ to represent those of the five ocular characters.

Let us now turn to the data and see how far we are able to some extent to answer the questions propounded above.

III. Problems and Results.

Problem 1. *For the population as a whole, for each character of hand and eye, which side is dominant and by how much?*

Table I gives the values of the means and standard deviations for each character of hand and eye based on the data (*a*) of the whole population, (*b*) of individuals showing dextrality, (*c*) of those showing sinistrality. From Table I it will be seen that, on the average, all ocular characters are a little superior in the right eye but that this superiority is insignificant except in the case of the two Refractions. For the manual characters the superiority of Grip is, in all cases, in favour of the right hand and is large, 4·53 lbs. This is approximately 1·3 lbs. higher than we obtained from Galton's data. The higher value for this character, obtained in the present investigation, may be due to the fact that the data were taken from College Students so that no young children and old people are included.

The steadiness and balancing tests are about $1\frac{1}{2}$ times better for the right hand than for the left. For the first pull and the difference between first and last pull the superiority of the right hand over the left is about $\frac{1}{2}$ lb. which is small if compared with the superiority of grip. In the case of the last pull, the left hand shows a slight superiority which, however, becomes insignificant when compared with its probable error.

Let us now study the fourth and sixth columns of Table I and compare the constants of each side. It is obvious that for all ocular characters the difference of means between both sides is very small, occurring only in the decimal places of the given units, and that the difference of variation between each corresponding pair is also small. These results appear to agree with those we obtained from Galton's data on Visual Acuity. The comparison of the right and left superiority (Columns 4 and 6) for three of the manual characters—namely the three pull tests—also shows that, although the values are not equal, there is only a slight difference between them. The right hand has, on the average, a superiority of 2·6 lbs. for the grip test and 1 unit for the steadiness and balancing tests. Their variabilities are a little greater on the right than on the left.

From all these results, it will be seen that, though most of the characters studied show a superiority of the right side the amount in more than half the cases is too small to be significant.

TABLE I.

Constants of Variates of Eye and Hand.

Characters	All Observations		Right side Superiority		Left side Superiority	
Ocular Laterality (Right minus Left)	Mean	S.D.	Mean	S.D.	Mean	S.D.
(1) Near Point	+0·0906 ± ·0397 ins. (414)	1·1970 ± ·0281 ins.	+1·2718 ± ·0796 ins. (103)	1·1977 ± ·0563 ins.	−1·3955 ± ·1261 ins. (67)	1·5302 ± ·0892 ins.
(2) Visual Acuity	+0·0045 ± ·0072 (419)	0·2188 ± ·0051	+0·1685 ± ·0107 (164)	0·2037 ± ·0076	−0·1384 ± ·0081 (186)	0·1646 ± ·0058
(3) Refraction-Horizontal Meridian	+0·1643 ± ·0266 D. (394)	0·7839 ± ·0188 D.	+0·6889 ± ·0336 D. (184)	0·6753 ± ·0237 D.	−0·6596 ± ·0461 D. (94)	0·6624 ± ·0326 D.
(4) Refraction-Vertical Meridian	+0·1437 ± ·0276 D. (395)	0·8143 ± ·0195 D.	+0·6753 ± ·0360 D. (184)	0·7240 ± ·0255 D.	−0·7849 ± ·0493 D. (86)	0·6783 ± ·0349 D.
(5) General Astigmatism	+0·0090 ± ·0197 D. (417)	0·5973 ± ·0140 D.	+0·6699 ± ·0449 D. (103)	0·6754 ± ·0317 D.	−0·5393 ± ·0258 D. (121)	0·4203 ± ·0182 D.
Manual Laterality (Right minus Left)						
(1) Grip	4·5316 ± ·1739 lbs. (459)	5·5228 ± ·1229 lbs.	+6·5374 ± ·1473 lbs. (361)	4·1488 ± ·1041 lbs.	−3·8889 ± ·2473 lbs. (72)	3·2184 ± ·1749 lbs.
(2) First Pull	+0·5811 ± ·1463 lbs. (456)	4·6299 ± ·1034 lbs.	+3·8996 ± ·1366 lbs. (239)	3·1315 ± ·0966 lbs.	−3·8555 ± ·1450 lbs. (173)	2·8274 ± ·1025 lbs.
(3) Last Pull	−0·1903 ± ·1443 lbs. (452)	4·5491 ± ·1020 lbs.	+3·5714 ± ·1225 lbs. (203)	2·5872 ± ·0866 lbs.	−4·0960 ± ·1563 lbs. (198)	3·2603 ± ·1105 lbs.
(4) Lack of Endurance of Pull	+0·7311 ± ·1751 lbs. (450)	5·5062 ± ·1238 lbs.	+4·8609 ± ·1664 lbs. (230)	3·7419 ± ·1177 lbs.	−4·3115 ± ·1577 lbs. (183)	3·1623 ± ·1115 lbs.
(5) Steadiness Test*	+1·5795 ± ·0876 (459)	2·7826 ± ·0619	+3·0803 ± ·0746 (299)	1·9121 ± ·0527	−2·0206 ± ·0948 (97)	1·3843 ± ·0670
(6) Balancing Test*	+1·8972 ± ·2233 (457)	7·0784 ± ·1579	+7·4810 ± ·1698 (237)	3·8754 ± ·1201	−6·6131 ± ·2083 (137)	3·6144 ± ·1473

* Units those of marking scores.

Therefore it is impossible to say that all the characters of hand and eye are always superior in one direction and that the amount of superiority, whatever its direction, is always the same proportion of the whole intensity ($\frac{1}{6}$) as van Biervliet* asserts.

Problems 2 and 3. *Is the hand or eye which is dominant for one manual or ocular character also dominant for the others? Is manual laterality absolute? Is ocular laterality absolute?*

(*a*) *On the relationships of percentage occurrences of the dominance of* (1) *Ocular*, (2) *Manual characters.*

It has been suggested by those who hold the theory of absolute laterality both for hand and eye that these characters follow the simple Mendelian ratio with dextrality as the dominant and sinistrality as the recessive character. This (according to them) means that three quarters of the population show dextrality and a quarter sinistrality while there are no ambilaterists at all. We have found that

TABLE II.

Percentages of Ocular and Manual Lateralities for all Observations.

a. Ocular Lateralities

Ocular characters	Sinistralists			Ambilateralists		Dextralists		
	Percentages	Ratios to 1st character	Diff. from Mendelian's 1 quarter	Percentages	Ratios to 1st character	Percentages	Ratios to 1st character	Diff. from Mendelian's 3 quarters
			%					%
Near Point	16·18±1·22 (67)	1·00	− 8·82	58·94±1·63 (244)	1·00	24·88±1·43 (103)	1·00	−50·18
Visual Acuity	44·39±1·64 (186)	2·74	+19·39	16·47±1·22 (69)	0·28	39·14±1·61 (164)	1·57	−35·86
Refraction in the Horizontal Meridian	23·86±1·45 (94)	1·47	− 1·14	29·44±1·55 (116)	0·50	46·70±1·70 (184)	1·88	−28·30
Refraction in the Vertical Meridian	21·77±1·39 (86)	1·35	− 3·23	31·65±1·58 (125)	0·54	46·58±1·68 (184)	1·87	−28·42
General Astigmatism	29·02±1·50 (121)	1·79	+ 4·02	46·28±1·65 (193)	0·79	24·70±1·43 (103)	0·99	−50·30

b. Manual Lateralities

Manual characters	Sinistralists			Ambilateralists		Dextralists		
	Percentages	Ratios to 1st character	Diff. from Mendelian's 1 quarter	Percentages	Ratios to 1st character	Percentages	Ratios to 1st character	Diff. from Mendelian's 3 quarters
			%					%
Grip	15·69±1·15 (72)	1·00	− 9·31	5·66±0·73 (26)	1·00	78·65±1·29 (361)	1·00	+ 3·65
First Pull	37·94±1·53 (173)	2·42	+12·94	9·65±0·93 (44)	1·70	52·41±1·58 (239)	0·67	−22·59
Last Pull	43·81±1·57 (198)	2·79	+18·81	11·28±1·00 (51)	1·99	44·91±1·58 (203)	0·57	−30·09
Lack of Endurance (First Pull−Last Pull)	40·67±1·56 (183)	2·59	+15·67	8·22±0·87 (37)	1·45	51·11±1·59 (230)	0·65	−23·89
Steadiness of Hand	21·13±1·29 (97)	1·35	− 3·87	13·73±1·08 (63)	2·43	65·14±1·50 (299)	0·83	− 9·86
Balancing Test	29·98±1·45 (137)	1·91	+ 4·98	18·16±1·22 (83)	3·21	51·86±1·58 (237)	0·66	−23·14

* "L'asymétrie sensorielle," *Bulletins de l'Académie royale de Belgique*, 3me série, Tom. XXXIV. pp. 326—366, Bruxelles, 1897.

Galton's data provided no confirmation of this hypothesis. Let us now see whether our present results support the theory of absolute laterality. Table II gives the percentage occurrence of each of the three main types of laterality for the characters studied in the total observations. It will be seen that the percentages of each type of laterality vary very considerably for different characters.

If each column be examined separately it will be found that the percentages do not correspond to those required by a supposed theory of Mendelian occurrence*. Only one of the sinistral characters nearly approaches a 25 per cent. occurrence, having regard to its probable error—this is the refraction taken from the horizontal meridian which gives a value of $23{\cdot}86 \pm 1{\cdot}45$ per cent. However, the dextralists for this character are only $46{\cdot}70 \pm 1{\cdot}70$ per cent. and the value for ambilateralists being $29{\cdot}44 \pm 1{\cdot}55$ per cent. shows them to be far from non-existent. None of the dextral percentages gives a result of 75 per cent. and all the figures for ambilateralists are significant. The smallest percentage of ambilateralists is for Grip, but although this is only $5{\cdot}66 \pm 0{\cdot}73$ per cent. it is definitely significant considering the value of the probable error. In order to compare all the characters used more clearly we have found the value of the deviation of the observed percentages from those required by the Mendelian ratio for both dextrality and sinistrality. There is a significant difference between all of them except for sinistrality in the case of horizontal refraction.

The divergence of the percentages for the various characters of each group can be tested in another way. The percentages of Near Point and of Grip can be supposed to be units for ocular and manual characters respectively and ratios of the percentages of other characters to that of the corresponding unit are calculated and set forth in the above table. It will be obvious how largely they differ from one another in amount. For instance, in the case of Near Point and Visual Acuity there are 2·74 times more sinistralists and 1·57 times more dextralists for the latter character but there are only 0·28 times more ambilateralists.

From the above results we may assert that neither manual nor ocular laterality is absolute. In other words, it does not follow that because the right hand or eye is dominant for one character it is necessarily dominant for another even though the two characters may be closely related.

We will now proceed to analyse our data more closely. So far we have only considered the laterality of each character, clubbing together all observations, and we may have confused the issues by only considering the total frequency of the laterality of one variate without studying the associate distribution of laterality of the two variates combined. For this reason, the percentage occurrences in nine categories of lateralism for any two characters of the two groups are set forth in Table III *a* and *b*. It involves twenty-five pairs of characters, ten for the ocular group and fifteen for the manual one. Each presents three kinds of percentages, namely, (1) those in round brackets are the actual percentages of the total population, (2) those in square brackets are percentages in the cells of

* See *American Breeders' Magazine*, Vol. II, pp. 19—113; *Journal of Genetics*, Vol. IV, pp. 67—81; *American Naturalist*, Vol. XLVII, pp. 734—738.

TABLE III.

Percentages Observed and Expected for each Category of (a) Ocular Lateralities, and (b) Manual Lateralities.

(*a*) Ocular Lateralities.

(1) Near Point and Visual Acuity.

Near Point	Visual Acuity									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	24 (5·80)	(6·88)	[13·64]	9 (2·17)	(2·77)	[12·68]	34 (8·21)	(6·53)	[20·36]	67 (16·18)
Ambilaterality	103 (24·88)	(25·05)	[58·52]	45 (10·87)	(10·12)	[63·38]	96 (23·19)	(23·77)	[57·49]	244 (58·94)
Dextrality	49 (11·83)	(10·58)	[27·84]	17 (4·11)	(4·26)	[23·94]	37 (8·94)	(10·04)	[22·15]	103 (24·88)
Totals	176 (42·51)			71 (17·15)			167 (40·34)			414 (100)

(2) Near Point and Refraction in "Horizontal" Meridian.

Near Point	Refraction in "Horizontal" Meridian									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	15 (3·87)	(3·43)	[15·96]	10 (2·58)	(4·25)	[8·62]	30 (7·73)	(6·50)	[16·85]	55 (14·18)
Ambilaterality	56 (14·42)	(14·73)	[59·57]	76 (19·59)	(18·18)	[65·52]	104 (26·81)	(27·91)	[58·43]	236 (60·82)
Dextrality	23 (5·93)	(6·06)	[24·47]	30 (7·73)	(7·47)	[25·86]	44 (11·34)	(11·47)	[24·72]	97 (25·00)
Totals	94 (24·22)			116 (29·90)			178 (45·88)			388 (100)

(3) Near Point and Refraction in "Vertical" Meridian.

Near Point	Refraction in "Vertical" Meridian									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	17 (4·37)	(3·33)	[21·25]	13 (3·34)	(5·37)	[10·08]	33 (8·48)	(7·49)	[18·33]	63 (16·19)
Ambilaterality	42 (10·79)	(12·32)	[52·50]	87 (22·37)	(19·86)	[67·44]	104 (26·74)	(27·72)	[57·78]	233 (59·90)
Dextrality	21 (5·41)	(4·92)	[26·25]	29 (7·45)	(7·93)	[22·48]	43 (11·05)	(11·06)	[23·89]	93 (23·91)
Totals	80 (20·57)			129 (33·16)			180 (46·27)			389 (100)

(4) Near Point and General Astigmatism.

Near Point	General Astigmatism									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	26 (6·33)	(4·40)	[22·03]	19 (4·62)	(7·20)	[9·84]	18 (4·38)	(3·73)	[18·00]	63 (15·33)
Ambilaterality	68 (16·54)	(17·12)	[57·63]	117 (28·47)	(27·99)	[60·62]	60 (14·60)	(14·50)	[60·00]	245 (59·61)
Dextrality	24 (5·84)	(7·19)	[20·34]	57 (13·87)	(11·77)	[29·54]	22 (5·35)	(6·10)	[22·00]	103 (25·06)
Totals	118 (28·71)			193 (46·96)			100 (24·33)			411 (100)

(5) Visual Acuity and Refraction in "Horizontal" Meridian.

Visual Acuity	Refraction in "Horizontal" Meridian									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	46 (11·71)	(9·79)	[50·55]	53 (13·48)	(12·79)	[44·54]	67 (17·05)	(19·66)	[36·61]	166 (42·24)
Ambilaterality	7 (1·78)	(4·00)	[7·69]	33 (8·40)	(5·24)	[27·73]	28 (7·12)	(8·06)	[15·30]	68 (17·30)
Dextrality	38 (9·67)	(9·37)	[41·76]	33 (8·40)	(12·25)	[27·73]	88 (22·39)	(18·84)	[48·09]	159 (40·46)
Totals	91 (23·16)			119 (30·28)			183 (46·56)			393 (100)

(6) Visual Acuity and Refraction in "Vertical" Meridian.

Visual Acuity	Refraction in "Vertical" Meridian									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	37 (9·37)	(9·21)	[44·05]	66 (16·70)	(14·02)	[51·56]	68 (17·22)	(20·06)	[37·16]	171 (43·29)
Ambilaterality	17 (4·30)	(3·71)	[20·24]	29 (7·35)	(5·67)	[22·66]	23 (5·82)	(8·09)	[12·57]	69 (17·47)
Dextrality	30 (7·59)	(8·34)	[35·71]	33 (8·36)	(12·72)	[25·78]	92 (23·29)	(18·18)	[50·27]	155 (39·24)
Totals	84 (21·26)			128 (32·41)			183 (46·33)			395 (100)

(7) Visual Acuity and General Astigmatism.

Visual Acuity	General Astigmatism									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	57 (13·68)	(12·53)	[47·11]	76 (18·22)	(19·98)	[39·38]	47 (11·27)	(10·66)	[45·63]	180 (43·17)
Ambilaterality	17 (4·07)	(4·94)	[14·05]	39 (9·35)	(7·87)	[20·21]	15 (3·60)	(4·21)	[14·56]	71 (17·02)
Dextrality	47 (11·27)	(11·55)	[38·84]	78 (18·71)	(18·43)	[40·41]	41 (9·83)	(9·83)	[39·81]	166 (39·81)
Totals	121 (29·02)			193 (46·28)			103 (24·70)			417 (100)

(8) Refractions in "Horizontal" and "Vertical" Meridians.

Refraction in "Horizontal" Meridian	Refraction in "Vertical" Meridian									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	47 (12·27)	(4·74)	[59·49]	18 (4·70)	(7·62)	[14·18]	23 (6·01)	(10·62)	[12·99]	88 (22·98)
Ambilaterality	14 (3·66)	(5·93)	[17·72]	76 (19·84)	(9·52)	[59·84]	20 (5·22)	(13·27)	[11·30]	110 (28·72)
Dextrality	18 (4·70)	(9·96)	[22·79]	33 (8·62)	(16·02)	[25·98]	134 (34·98)	(22·32)	[75·71]	185 (48·30)
Totals	79 (20·63)			127 (33·16)			177 (46·21)			383 (100)

(9) Refraction in "Horizontal" Meridian and General Astigmatism.

Refraction in "Horizontal" Meridian	General Astigmatism									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	46 (11·68)	(6·73)	[41·44]	27 (6·85)	(11·26)	[14·51]	21 (5·33)	(5·87)	[21·65]	94 (23·86)
Ambilaterality	28 (7·10)	(8·29)	[25·23]	77 (19·55)	(13·90)	[41·40]	11 (2·79)	(7·25)	[11·34]	116 (29·44)
Dextrality	37 (9·39)	(13·15)	[33·33]	82 (20·81)	(22·05)	[44·09]	65 (16·50)	(11·50)	[67·01]	184 (46·70)
Totals	111 (28·17)			186 (47·21)			97 (24·62)			394 (100)

(10) Refraction in "Vertical" Meridian and General Astigmatism.

Refraction in "Vertical" Meridian	General Astigmatism									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	31 (7·86)	(6·07)	[28·18]	28 (7·08)	(10·47)	[14·74]	27 (6·83)	(5·23)	[28·42]	86 (21·77)
Ambilaterality	19 (4·81)	(8·81)	[17·27]	81 (20·51)	(15·22)	[42·63]	25 (6·33)	(7·62)	[26·32]	125 (31·65)
Dextrality	60 (15·18)	(12·97)	[54·55]	81 (20·51)	(22·41)	[42·63]	43 (10·89)	(11·20)	[45·26]	184 (46·58)
Totals	110 (27·85)			190 (48·10)			95 (24·05)			395 (100)

(*b*) Manual Lateralities.

(1) Grip and First Pull.

Grip	First Pull									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	36 (7·93)	(5·93)	[20·93]	7 (1·54)	(1·51)	[15·91]	28 (6·17)	(8·20)	[11·77]	71 (15·64)
Ambilaterality	10 (2·21)	(2·17)	[5·81]	1 (0·22)	(0·56)	[02·27]	15 (3·30)	(3·00)	[6·30]	26 (5·73)
Dextrality	126 (27·75)	(29·79)	[73·26]	36 (7·93)	(7·62)	[81·82]	195 (42·95)	(41·22)	[81·93]	357 (78·63)
Totals	172 (37·89)			44 (9·69)			238 (52·42)			454 (100)

(2) Grip and Last Pull.

Grip	Last Pull									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	37 (8·23)	(6·91)	[18·78]	5 (1·11)	(1·75)	[10·00]	29 (6·44)	(7·12)	[14·29]	71 (15·78)
Ambilaterality	16 (3·55)	(2·43)	[8·12]	0 (0·00)	(0·62)	[0·00]	9 (2·00)	(2·50)	[4·43]	25 (5·55)
Dextrality	144 (32·00)	(34·44)	[73·10]	45 (10·00)	(8·74)	[90·00]	165 (36·67)	(35·49)	[81·28]	354 (78·67)
Totals	197 (43·78)			50 (11·11)			203 (45·11)			450 (100)

(3) Grip and Lack of Endurance in Pulling (First Pull – Last Pull).

Grip	Lack of Endurance in Pulling									Totals %
	Sinistrality			Ambilateraliy			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	32 (7·16)	(6·35)	[17·39]	6 (1·34)	(1·28)	[16·22]	31 (6·94)	(7·81)	[13·72]	69 (15·44)
Ambilateraliy	8 (1·79)	(2·30)	[4·35]	0 (0·00)	(0·46)	[00·00]	17 (3·80)	(2·83)	[7·52]	25 (5·59)
Dextrality	144 (32·21)	(32·51)	[78·26]	31 (6·94)	(6·54)	[83·78]	178 (39·82)	(39·92)	[78·76]	353 (78·97)
Totals	184 (41·16)			37 (8·28)			226 (50·56)			447 (100)

(4) Grip and Steadiness of Hand Test.

Grip	Steadiness of Hand Test									Totals %
	Sinistrality			Ambilateraliy			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	19 (4·14)	(3·31)	[19·59]	7 (1·53)	(2·15)	[11·11]	46 (10·02)	(10·23)	[15·38]	72 (15·69)
Ambilateraliy	5 (1·09)	(1·20)	[5·15]	3 (0·65)	(0·78)	[4·76]	18 (3·92)	(3·68)	[6·02]	26 (5·66)
Dextrality	73 (15·90)	(16·62)	[75·26]	53 (11·55)	(10·80)	[84·13]	235 (51·20)	(51·23)	[78·60]	361 (78·65)
Totals	97 (21·13)			63 (13·73)			299 (65·14)			459 (100)

(5) Grip and Balancing Test.

Grip	Balancing Test									Totals %
	Sinistrality			Ambilateraliy			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	29 (6·39)	(4·68)	[21·97]	13 (2·86)	(2·94)	[15·66]	31 (6·83)	(8·46)	[12·97]	73 (16·08)
Ambilateraliy	6 (1·33)	(1·60)	[4·55]	5 (1·10)	(1·01)	[6·02]	14 (3·08)	(2·90)	[5·86]	25 (5·51)
Dextrality	97 (21·36)	(22·80)	[73·48]	65 (14·32)	(14·33)	[78·32]	194 (42·73)	(41·28)	[81·17]	356 (78·41)
Totals	132 (29·08)			83 (18·28)			239 (52·64)			454 (100)

(6) First Pull and Last Pull.

First Pull	Last Pull									Totals %
	Sinistrality			Ambilateraliy			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	94 (20·80)	(16·57)	[47·47]	18 (3·98)	(4·27)	[35·29]	59 (13·05)	(16·99)	[29·07]	171 (37·83)
Ambilateraliy	23 (5·09)	(4·37)	[11·62]	5 (1·11)	(1·12)	[9·81]	17 (3·76)	(4·47)	[8·37]	45 (9·96)
Dextrality	81 (17·92)	(22·87)	[40·91]	28 (6·19)	(5·89)	[54·90]	127 (28·10)	(23·45)	[62·56]	236 (52·21)
Totals	198 (43·81)			51 (11·28)			203 (44·91)			452 (100)

(7) First Pull and Lack of Endurance in Pulling.

First Pull	Lack of Endurance in Pulling									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	113 (25·11)	(15·36)	[61·75]	14 (3·11)	(3·11)	[37·84]	43 (9·56)	(19·31)	[18·70]	170 (37·78)
Ambilaterality	18 (4·01)	(3·98)	[9·84]	5 (1·11)	(0·80)	[13·51]	21 (4·66)	(5·00)	[9·13]	44 (9·78)
Dextrality	52 (11·55)	(21·33)	[28·41]	18 (4·00)	(4·31)	[48·65]	166 (36·89)	(26·80)	[72·17]	236 (52·44)
Totals	183 (40·67)			37 (8·22)			230 (51·11)			450 (100)

(8) First Pull and Steadiness of Hand Test.

First Pull	Steadiness of Hand Test									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp		
Sinistrality	41 (8·99)	(8·07)	[42·27]	26 (5·70)	(5·16)	[41·94]	106 (23·25)	(24·71)	[35·69]	173 (37·94)
Ambilaterality	10 (2·19)	(2·06)	[10·31]	2 (0·44)	(1·31)	[3·22]	32 (7·02)	(6·28)	[10·77]	44 (9·65)
Dextrality	46 (10·09)	(11·14)	[47·42]	34 (7·46)	(7·13)	[54·84]	159 (34·86)	(34·14)	[53·54]	239 (52·41)
Totals	97 (21·27)			62 (13·60)			297 (65·13)			456 (100)

(9) First Pull and Balancing Test.

First Pull	Balancing Test									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	58 (12·86)	(11·35)	[42·96]	28 (6·21)	(6·89)	[34·15]	85 (18·85)	(19·68)	[36·33]	171 (37·92)
Ambilaterality	13 (2·88)	(2·92)	[9·63]	9 (2·00)	(1·77)	[10·98]	22 (4·87)	(5·06)	[9·40]	44 (9·75)
Dextrality	64 (14·19)	(15·66)	[47·41]	45 (9·97)	(9·52)	[54·87]	127 (28·17)	(27·15)	[54·27]	236 (52·33)
Totals	135 (29·93)			82 (18·18)			234 (51·89)			451 (100)

(10) Last Pull and Lack of Endurance in Pulling.

Last Pull	Lack of Endurance in Pulling									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	37 (8·23)	(17·82)	[20·11]	14 (3·11)	(3·58)	[37·84]	145 (32·22)	(22·16)	[63·32]	196 (43·56)
Ambilaterality	18 (4·00)	(4·54)	[9·78]	5 (1·11)	(0·91)	[13·51]	27 (6·00)	(5·66)	[11·79]	50 (11·11)
Dextrality	129 (28·66)	(18·53)	[70·11]	18 (4·00)	(3·73)	[48·65]	57 (12·67)	(23·07)	[24·89]	204 (45·33)
Totals	184 (40·89)			37 (8·22)			229 (50·89)			450 (100)

(11) Last Pull and Steadiness of Hand Test.

Last Pull	Steadiness of Hand Test									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	48 (10·64)	(9·25)	[50·53]	29 (6·43)	(6·04)	[46·77]	121 (26·83)	(28·61)	[41·16]	198 (43·90)
Ambilaterality	9 (2·00)	(2·33)	[9·47]	4 (0·89)	(1·52)	[6·46]	37 (8·20)	(7·24)	[12·58]	50 (11·09)
Dextrality	38 (8·42)	(9·48)	[40·00]	29 (6·43)	(6·19)	[46·77]	136 (30·16)	(29·34)	[46·26]	203 (45·01)
Totals	95 (21·06)			62 (13·75)			294 (65·19)			451 (100)

(12) Last Pull and Balancing Test.

Last Pull	Balancing Test									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	61 (13·74)	(13·19)	[45·52]	39 (8·78)	(7·87)	[48·75]	94 (21·17)	(22·63)	[40·87]	194 (43·69)
Ambilaterality	10 (2·25)	(3·40)	[7·46]	12 (2·70)	(2·03)	[15·00]	28 (6·31)	(5·83)	[12·17]	50 (11·26)
Dextrality	63 (14·19)	(13·59)	[47·02]	29 (6·54)	(8·12)	[36·25]	108 (24·32)	(23·34)	[46·96]	200 (45·05)
Totals	134 (30·18)			80 (18·02)			230 (51·80)			444 (100)

(13) Steadiness of Hand and Lack of Endurance in Pulling.

Steadiness of Hand	Lack of Endurance in Pulling									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	43 (9·56)	(8·68)	[22·99]	10 (2·22)	(1·72)	[27·03]	41 (9·11)	(10·49)	[18·14]	94 (20·89)
Ambilaterality	26 (5·77)	(5·54)	[13·91]	4 (0·89)	(1·09)	[10·81]	30 (6·67)	(6·70)	[13·27]	60 (13·33)
Dextrality	118 (26·23)	(27·34)	[63·10]	23 (5·11)	(5·41)	[62·16]	155 (34·44)	(33·03)	[68·59]	296 (65·78)
Totals	187 (41·56)			37 (8·22)			226 (50·22)			450 (100)

(14) Balancing Test and Lack of Endurance in Pulling.

Balancing Test	Lack of Endurance in Pulling									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	60 (13·51)	(12·09)	[32·97]	9 (2·04)	(2·46)	[24·33]	62 (13·96)	(14·96)	[27·55]	131 (29·51)
Ambilaterality	31 (6·98)	(7·66)	[17·03]	5 (1·12)	(1·56)	[13·51]	47 (10·59)	(9·47)	[20·89]	83 (18·69)
Dextrality	91 (20·50)	(21·24)	[50·00]	23 (5·17)	(4·31)	[62·16]	116 (26·13)	(26·25)	[51·56]	230 (51·80)
Totals	182 (40·99)			37 (8·33)			225 (50·68)			444 (100)

(15) Steadiness of Hand and Balancing Test.

Steadiness of Hand	Balancing Test									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	21 (4·60)	(6·30)	[15·33]	19 (4·16)	(3·82)	[22·89]	56 (12·25)	(10·89)	[23·63]	96 (21·01)
Ambilaterality	24 (5·25)	(4·07)	[17·52]	12 (2·62)	(2·46)	[14·46]	26 (5·70)	(7·04)	[10·97]	62 (13·57)
Dextrality	92 (20·13)	(19·61)	[67·15]	52 (11·38)	(11·88)	[62·65]	155 (33·91)	(33·93)	[65·40]	299 (65·42)
Totals	137 (29·98)			83 (18·16)			237 (51·86)			457 (100)

the numbers at the foot of the column, and (3) those under the columns headed "expected" give the percentages which might be anticipated if there were no relation whatever between any two characters. These pairs of values may be tested by such a criterion as this: If little relation exists for the given traits, the percentages of laterality of one character for each type of laterality of another should be sensibly the same as in the corresponding totals column, and also the difference in each category between the observed percentage and the expected should be very slight. On the contrary in case a high positive relation exists between the two given traits, the observed value in the categories of the same laterality (or column percentage in square brackets) will be larger than the expected one (or marginal percentage). Those in the categories of opposite laterality will be much less than the corresponding independent values, and the higher the correlation, the larger the difference. If the correlation be negative and high, the difference between the percentages should be in the reverse order. By the criterion just mentioned we may test each pair of the percentages in the above table. A quantitative relation between any two characters may be roughly predicted from the knowledge of the nature and the degree of those differences. From our material it will be found that in the Manual Group there are two cases, the dominance of First Pull (7) and Last Pull (10) with that of Lack of Endurance, where there are high correlations but with opposite signs; and in another two, the dominance of First Pull with dominance of Grip (1) and of First Pull with Last Pull (6), with moderate but significant correlations. Similarly in the Visual Group we have one case, the dominance of the two Refractions, with high positive correlation, and two cases, the dominance of Visual Acuity with that of the two Refractions, with moderate correlation. Besides these, the differences between the percentages for any other pair of characters are either not so appreciable or are definitely insignificant. These results show a good degree of agreement with the quantitative values of association which will be discussed later.

For the purpose of throwing further light on this material we have classified the actual observed percentages of all characters of the two groups into three main classes which will be described as follows:

(1) Same laterality involving individuals of same dextrality, same sinistrality and same ambilaterality.

(2) Opposite laterality involving those who are dextral in one variate but sinistral in the other of the same group.

(3) Mixed laterality involving those who are dextral or sinistral in one variate but ambilateral in the other.

Table IV provides the classified percentages in each type of laterality for all the characters of both groups. The mean percentages and standard deviations are also set down at the foot of each column respectively. When this table has been carefully examined the reader will conclude that the mean percentage of the same laterality in the case of both hand and eye is always the largest but that its value, which is less than 50 per cent. of the whole population, cannot be said to form a law of that population. Even in such a highly related pair of two traits as the two Refractions, the value of the same laterality still does not exceed 70 per cent. of the whole. Another high value of the same laterality is found in the manual group in the case of First Pull and Lack of Endurance in pulling. But still there are 38 per cent. of individuals who do not follow the supposed rule. The mean percentage of opposite laterality for each group is comparatively low in its amount but the value for each group (approximately one-third in the manual group and one-sixth in the visual group) is amply sufficient to prove the fallibility of the proposed rule. Also if the numbers for mixed laterality be added we find that the individuals with the same laterality are actually in a minority.

Further, if we consider the percentages of like laterality in any two characters in each column, it will still be found that none of the individual values for the common dextrality or for the common sinistrality is identical with the Mendelian ratio with the single exception of the common sinistrality for First Pull and Lack of Endurance (25·11 per cent.); and this is of no significance because Mendelian theory would demand 75 per cent. of the corresponding dextrality.

The numbers of common ambilateralists for two characters are really very few, especially in the manual group, but the reader must remember that this does not mean that all the ambilateralists for any one character are non-existent. Take an extreme case, that of Grip and Last Pull for illustration. The number of common ambilateralists is actually zero but if the value of laterality of each character be studied separately we should have 11·11 per cent. of ambilateralists for Last Pull and 5·55 per cent. for Grip (see Table III *b* (2)). The same is true in the case of Grip and Lack of Endurance. Among the mixed lateralities we find that the cases of ambilaterality combined with dextrality are significantly higher than those combined with sinistrality particularly in the visual group.

The standard deviations for each type of laterality in both groups are all significantly large, i.e. there is a great variation of percentages from sample to sample against the rule of absolute superiority.

From the consideration of these percentage occurrences we realise thoroughly that in each category of lateralism the relation of any two variates, either of the

TABLE IV. *(b) Manual Characters.*

TABLE IV.

Percentages of Types of Laterality for (a) Ocular Characters, (b) Manual Characters.

(*a*) Ocular Laterality.

Pairs of Characters (Observation)	(I) Same Laterality				(II) Opposite Laterality	(III) Mixed Laterality		
	Sinistrality	Ambi-laterality	Dextrality	Totals	Sinistrality and Dextrality	Sinistrality and Ambi-laterality	Dextrality and Ambi-laterality	Totals
(1) Near Point and Visual Acuity ... ... (414)	5·80	10·87	8·94	25·61	20·04	27·05	27·30	54·35
(2) Near Point and Refraction ("Horizontal" Meridian) ... ... (388)	3·87	19·59	11·34	34·80	13·66	17·00	34·54	51·54
(3) Near Point and Refraction ("Vertical" Meridian) (389)	4·37	22·37	11·05	37·79	13·89	14·13	34·19	48·32
(4) Near Point and General Astigmatism ... ... (411)	6·33	28·47	5·35	40·15	10·22	21·16	28·47	49·63
(5) Visual Acuity and Refraction "Horizontal" ... (393)	11·71	8·40	22·39	42·50	26·72	15·26	15·52	30·78
(6) Visual Acuity and Refraction "Vertical" ... (395)	9·37	7·35	23·29	40·01	24·81	21·00	14·18	35·18
(7) Visual Acuity and General Astigmatism ... (417)	13·68	9·35	9·83	32·86	22·54	22·29	22·31	44·60
(8) Refractions in the two Meridians ... ... (383)	12·27	19·84	34·98	67·09	10·71	8·36	13·84	22·20
(9) Refraction "Horizontal" and General Astigmatism (394)	11·68	19·55	16·50	47·73	14·72	13·95	23·60	37·55
(10) Refraction "Vertical" and General Astigmatism (395)	7·86	20·51	10·89	39·26	22·01	11·89	26·84	38·73
Mean	8·69±·72	16·63±1·44	15·46±1·82	40·78±2·22	17·93±1·21	17·21±1·13	24·08±1·55	41·29±2·06
Standard Deviation	3·37±·51	6·75±1·02	8·51±1·28	10·41±1·57	5·69±·86	5·32±·80	7·27±1·00	9·68±1·46

(*b*) Manual Laterality.

Pairs of Characters (Observation)	(I) Same Laterality				(II) Opposite Laterality	(III) Mixed Laterality		
	Sinistrality	Ambi-laterality	Dextrality	Totals	Sinistrality and Dextrality	Sinistrality and Ambi-laterality	Dextrality and Ambi-laterality	Totals
(1) Grip and First Pull ... ... (454)	7·93	0·22	42·95	51·10	33·92	3·75	11·23	14·98
(2) Grip and Last Pull ... ... (450)	8·23	0·00	36·67	44·90	38·44	4·66	12·00	16·66
(3) Grip and Steadiness Test ... ... (459)	4·14	0·65	51·20	55·99	25·92	2·62	15·47	18·09
(4) Grip and Balancing Test ... ... (454)	6·39	1·10	42·73	50·22	28·19	4·19	17·40	21·59
(5) Grip and Lack of Endurance of Pull ... ... (447)	7·16	0·00	39·82	46·98	39·15	3·13	10·74	13·87
(6) First Pull and Last Pull ... ... (452)	20·80	1·11	28·10	50·01	30·97	9·07	9·95	19·02
(7) First Pull and Steadiness Test ... ... (456)	8·99	0·44	34·86	44·29	33·34	7·89	14·48	22·37
(8) First Pull and Balancing Test ... ... (451)	12·86	2·00	28·17	43·03	33·04	9·09	14·84	23·93
(9) First Pull and Lack of Endurance of Pull ... (450)	25·11	1·11	36·89	63·11	21·11	7·12	8·66	15·78
(10) Last Pull and Steadiness Test ... ... (451)	10·64	0·89	30·16	41·69	35·25	8·43	14·63	23·06
(11) Last Pull and Balancing Test ... ... (444)	13·74	2·70	24·32	40·76	35·36	11·03	12·85	23·88
(12) Last Pull and Lack of Endurance of Pull ... (450)	8·23	1·11	12·67	22·01	60·88	7·11	10·00	17·11
(13) Steadiness Test and Lack of Endurance of Pull (450)	9·56	0·89	34·44	44·89	35·34	7·99	11·78	19·77
(14) Balancing Test and Lack of Endurance of Pull (450)	13·51	1·12	26·13	40·76	34·46	9·02	15·76	24·78
(15) Steadiness Test and Balancing Test ... ... (457)	4·60	2·62	33·91	41·13	32·38	9·41	17·08	26·49
Mean	10·79±·97	1·06±·14	33·54±1·55	45·39±1·52	34·52±1·45	6·97±·44	13·12±·47	20·09±·67
Standard Deviation	5·58±·69	0·80±·10	8·90±1·10	8·70±1·07	8·35±1·03	2·53±·31	2·68±·33	3·82±·47

manual group or the ocular group, is very largely independent and that the degree of laterality can in no way be treated as absolutely fixed for all characters.

(*b*) *On the Quantitative Interrelations of the Dominance of the Ocular Characters.*

We have thus far investigated the relationships of laterality in various manual and ocular characters by comparing the percentage occurrences of the observed value with those of the independent values for each type of laterality. We will now consider their quantitative associations by deducing the correlation values, i.e. those for r and η. If the theory of absolute laterality has any truth in it we should expect to find that one difference in visual character was closely associated with all the other differences in visual characters studied. The same should also be true for the manual characters. Let us first investigate the interrelations of the differences of the visual characters.

(1) Difference of Near Point (B_1) with those of the other four visual characters (B_2, B_3, B_4 and B_5).

The values of the product moment coefficients for the dominance of Near Point and that of the other characters are stated below. The number of subjects tested is given in brackets.

Dominance of Near Point and Visual Acuity: $r_{B_1B_2} = \cdot1122 \pm \cdot0327$ (414).

Dominance of Near Point and Horizontal Refraction: $r_{B_1B_3} = \cdot1028 \pm \cdot0339$ (388).

Dominance of Near Point and Vertical Refraction: $r_{B_1B_4} = \cdot0069 \pm \cdot0342$ (389).

Dominance of Near Point and General Astigmatism: $r_{B_1B_5} = \cdot0206 \pm \cdot0333$ (411).

It will be seen that all these coefficients have a positive value but that they are of very low intensity. The only one of the above coefficients which may probably be regarded as significant is $r_{B_1B_2}$ correlating the lateral difference of Near Point and Visual Acuity—this has a value 3·5 times its probable error. It means that an increase in the dominance of the Near Point, i.e. of near vision, is associated to a very small extent with increase in the dominance of distant vision. The coefficient measuring the relation of Near Point and Horizontal Refraction is just three times its probable error which probably indicates a slight relation. In other words, the eye dominant for near vision will be very slightly dominant for horizontal refraction.

The correlation ratios are as follows*:

$$\begin{cases} B_1 \text{ on } B_2: & \eta^2 = \cdot039241, & \overline{\eta}^2 = \cdot028986 \pm \cdot007856, \\ B_2 \text{ on } B_1: & \eta_2 = \cdot023034, & \overline{\eta}^2 = \cdot012077 \pm \cdot005115, \end{cases}$$

$$\begin{cases} B_1 \text{ on } B_3: & \eta^2 = \cdot030875, & \overline{\eta}^2 = \cdot018041 \pm \cdot006437, \\ B_3 \text{ on } B_1: & \eta^2 = \cdot029833, & \overline{\eta}^2 = \cdot012887 \pm \cdot005455, \end{cases}$$

$$\begin{cases} B_1 \text{ on } B_4: & \eta^2 = \cdot034814, & \overline{\eta}^2 = \cdot020566 \pm \cdot006854, \\ B_4 \text{ on } B_1: & \eta^2 = \cdot010406, & \overline{\eta}^2 = \cdot012853 \pm \cdot005440, \end{cases}$$

$$\begin{cases} B_1 \text{ on } B_5: & \eta^2 = \cdot029790, & \overline{\eta}^2 = \cdot014599 \pm \cdot005636, \\ B_5 \text{ on } B_1: & \eta^2 = \cdot017959, & \overline{\eta}^2 = \cdot012165 \pm \cdot005152. \end{cases}$$

* $\overline{\eta}^2 = (\kappa - 1)/N \pm \cdot67449\sqrt{2\overline{\eta}^2(1-\overline{\eta}^2)/(N+1)}$.

In seven out of the eight cases the value of η^2 is insignificant when compared with the corresponding value of $\overline{\eta}^2$. In the fourth example shown the difference between η^2 and $\overline{\eta}^2$ is about 3·1 times the probable error, thus lying just on the border line of significance.

Since the above results show no correlations of lateralities of any importance it has not been thought necessary to give correlation tables or diagrams.

(2) Difference of Visual Acuity with those of the two Refractions and of General Astigmatism.

Next, we shall consider the influence, if any, of dominance of Visual Acuity on the dominance of both the Refractions and General Astigmatism. The correlation constants deduced from the figures of Tables V and VI are given below.

Product moment correlation coefficients

$r_{B_2B_3} = +\cdot2056 \pm \cdot0326\ (393)$,

Correlation ratios

$\begin{cases} B_2 \text{ on } B_3: & \eta = \cdot2487, \quad \eta^2 = \cdot061832, \quad \overline{\eta}^2 = \cdot020356 \pm \cdot006786, \\ B_3 \text{ on } B_2: & \eta = \cdot2331, \quad \eta^2 = \cdot054325, \quad \overline{\eta}^2 = \cdot030534 \pm \cdot008268, \end{cases}$

$r_{B_2B_4} = +\cdot2062 \pm \cdot0325\ (395)$,

$\begin{cases} B_2 \text{ on } B_4: & \eta = \cdot2396, \quad \eta^2 = \cdot057399, \quad \overline{\eta}^2 = \cdot020253 \pm \cdot006751, \\ B_4 \text{ on } B_2: & \eta = \cdot2347, \quad \eta^2 = \cdot055094, \quad \overline{\eta}^2 = \cdot030380 \pm \cdot008226. \end{cases}$

Clearly the values of $r_{B_2B_3}$ and $r_{B_2B_4}$ show the same order of intensity and are considerable enough to be significant. This suggests that improvements of the refractions in dioptres are to some extent associated with increase of dominance of Visual Acuity. We have also four correlation ratios of approximately the same value, a little higher than r, and certainly significant.

Diagrams I and II show the two pairs of regression lines, from which we can see that the regressions of B_3 and B_4 on B_2 are nearly linear and can, accordingly, be fitted with linear equations.

(*a*) Probable $B_3 = \cdot72467 B_2 + \cdot19974$.

(*b*) Probable $B_4 = \cdot74429 B_2 + \cdot15740$.

The regressions of B_2 on B_3 and B_4 are not so simple, they show some degree of curvature and have been fitted with cubic equations.

(*c*) Probable $B_2 = -\cdot012216 + \cdot0753281 B_3 - \cdot00963851 B_3^2 - \cdot000372511 B_3^3$.

(*d*) Probable $B_2 = -\cdot003739 + \cdot0506592 B_4 - \cdot00904642 B_4^2 + \cdot000507456 B_4^3$.

The value of zero in B_3 or B_4 corresponds to $-\cdot01222$ of B_2 in equation (*c*) and to $-\cdot00374$ of B_2 in equation (*d*) which differ very little in amount. In other words, the chance of the dominance of Visual Acuity gives a somewhat similar chance of the dominance of two kinds of Refractions.

By the usual methods the correlation constants for the dominance of Visual Acuity and General Astigmatism are as follows:

$r_{B_2B_5} = \cdot0612 \pm \cdot0330\ (417)$, $\begin{cases} B_2 \text{ on } B_5: & \eta^2 = \cdot018334, \quad \overline{\eta}^2 = \cdot014388 \pm \cdot005556, \\ B_5 \text{ on } B_2: & \eta^2 = \cdot008572, \quad \overline{\eta}^2 = \cdot028777 \pm \cdot007799. \end{cases}$

TABLE V.

Laterality of Visual Acuity and Refraction in the Horizontal Meridian.

Difference of Refraction in the Horizontal Meridian in Dioptres. (Central Values.) Right minus Left.

Difference of Visual Acuity measured by Snellen's Type reduced to decimals. (Central Values.) Right minus Left.

Difference of Visual Acuity measured by Snellen's Type reduced to decimals. (Central Values.) Right minus Left.	+3·50	..	+2·75	..	+2·00	+1·75	+1·50	+1·25	+1·00	+0·75	+0·50	+0·25	0	−0·25	−0·50	−0·75	−1·00	−1·25	−1·50	−1·75	−2·00	−2·25	−2·50	−2·75	−3·00	−3·25	−3·50	..	−4·25	−4·50	
−1·11	—	..	—	..	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	1
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
−·96	—	..	—	..	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	1
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
−·84	—	..	—	..	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	1
−·81	—	..	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	0
−·78	—	..	—	..	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	1
−·75	—	..	1	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	1
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
−·66	—	..	—	..	—	—	—	—	—	1	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—	..	—	—	3
−·63	—	..	—	..	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	1
−·60	—	..	—	..	—	—	—	—	—	—	—	—	1	—	—	—	—	1	—	—	—	—	—	—	—	—	—	..	—	—	2
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
−·51	—	..	—	..	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	1
−·48	—	..	—	..	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	1
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
−·39	—	..	—	..	—	—	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	2
−·36	—	..	—	..	—	—	—	—	—	—	1	—	1	—	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	3
−·33	—	..	—	..	—	—	—	—	—	—	—	1	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	3
−·30	—	..	—	..	—	—	1	—	1	—	—	—	1	—	1	1	—	—	—	—	—	—	—	—	—	—	—	..	—	—	5
−·27	—	..	—	..	—	—	—	—	—	1	—	—	—	1	1	1	—	—	—	—	—	—	—	—	—	—	—	..	—	—	4
−·24	—	..	—	..	—	—	—	—	—	—	1	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	2
−·21	—	..	—	..	—	—	—	—	—	—	—	—	3	—	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	4
−·18	—	..	—	..	—	—	—	1	1	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—	1	—	..	—	—	5
−·15	—	..	—	..	—	—	—	—	—	—	1	6	2	2	1	2	—	—	—	—	—	—	—	—	—	—	—	..	—	—	14
−·12	—	..	—	..	—	—	—	—	—	—	3	2	5	3	3	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	16
−·09	—	..	—	..	—	—	—	—	2	1	1	—	8	2	5	1	—	—	—	—	—	—	—	—	—	—	—	..	—	—	20
−·06	—	..	—	..	—	1	—	—	—	2	2	2	12	6	7	1	1	—	—	—	—	1	—	—	—	—	—	..	—	—	35
−·03	—	..	—	..	—	—	—	1	1	—	3	3	18	5	4	3	1	—	—	—	—	—	—	—	—	—	—	..	1	—	40
0	—	..	—	..	1	—	1	—	1	2	1	1	33	10	6	2	1	—	3	—	1	—	1	—	3	—	—	..	1	—	68
·03	1	..	—	..	1	—	—	—	—	1	3	9	8	5	2	5	1	—	—	1	1	—	—	—	—	—	—	..	—	—	38
·06	—	..	—	..	—	—	—	—	—	—	—	3	2	6	7	10	2	—	—	—	—	—	—	—	—	—	—	..	—	—	30
·09	—	..	—	..	—	—	—	—	—	—	—	—	4	7	5	2	2	—	1	—	—	—	—	—	1	—	—	..	—	—	22
·12	—	..	—	..	—	—	—	—	—	—	1	—	2	4	3	4	2	—	—	—	—	—	—	—	—	—	—	..	—	—	16
·15	—	..	—	..	—	—	—	—	—	—	—	—	2	2	2	2	1	—	1	—	—	—	—	—	—	—	—	..	—	—	10
·18	—	..	—	..	—	—	—	—	1	1	—	—	—	1	3	2	1	—	—	—	—	—	—	—	—	—	—	..	—	—	9
·21	—	..	—	..	—	—	—	—	—	—	—	1	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	2
·24	—	..	—	..	—	—	—	—	—	—	1	1	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	..	—	—	3
·27	—	..	—	..	—	—	—	—	—	—	—	1	1	1	2	1	1	—	—	—	—	—	—	—	—	—	1	..	—	—	8
·30	—	..	—	..	—	—	—	—	—	—	—	—	3	1	2	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	6
·33	—	..	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	—	—	—	—	—	1	—	—	..	—	—	3
·36	—	..	—	..	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	..	—	—	1
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
·48	—	..	—	..	—	—	—	—	—	—	1	—	—	—	1	—	—	1	—	—	—	—	—	—	—	—	—	..	—	—	3
·51	—	..	—	..	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	1
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
·69	—	..	—	..	—	—	—	—	—	—	—	—	1	—	—	1	—	—	—	—	—	—	—	—	—	—	—	..	—	—	2
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
·84	—	..	—	..	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	..	—	—	1
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
·96	—	..	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	..	—	—	2
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1·05	—	..	—	..	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	..	—	—	1
1·08	—	..	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	1
	1	..	1	..	2	1	3	2	7	11	22	41	119	61	61	27	7	6	6	1	3	0	1	1	5	1	1	..	1	1	393

Difference of Visual Acuity measured by Snellen's Type reduced to decimals. (Central Values.) Right minus Left.

TABLE VI.

Laterality of Visual Acuity and Refraction in the Vertical Meridian.

Difference of Refraction in the Vertical Meridian in Dioptres. (Central Values.) Right minus Left.

Difference of Visual Acuity measured by Snellen's Type reduced to decimals. (Central Values.) Right minus Left.

Difference of Visual Acuity	+3·00	+2·75	+2·50	+2·25	+2·00	+1·75	+1·50	+1·25	+1·00	+0·75	+0·50	+0·25	0	−0·25	−0·50	−0·75	−1·00	−1·25	−1·50	−1·75	−2·00	−2·25	−2·50	−2·75	−3·00	−3·25	−3·50	..	−4·50	
−1·11	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	1
−·96	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	—	..
−·78	—	1	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	1
−·66	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	2	..	..	..	..	..	..	..	..	..	..	..	—	2
−·63	—	—	—	—	—	—	1	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	3
−·60	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	1	—	—	..	—	1
−·51	..	..	..	..	..	..	..	..	..	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	—	2
−·48	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	1
−·45	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	1
−·42	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	..	—	0
−·39	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	1
−·36	—	—	—	—	—	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	2
−·33	—	—	—	—	—	—	—	—	—	—	—	—	3	—	—	—	—	1	—	—	—	—	—	—	—	—	—	..	—	2
−·30	—	—	1	—	—	—	—	—	1	—	—	—	2	2	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	4
−·27	—	—	1	—	—	—	—	—	—	—	—	—	2	—	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	6
−·24	—	—	—	—	—	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	4
−·21	—	—	—	—	—	—	—	—	—	—	—	1	1	—	2	1	—	—	—	—	—	—	—	—	—	—	—	..	—	2
−·18	—	—	—	—	—	—	—	1	1	1	—	—	—	—	1	1	—	—	—	—	1	—	—	—	—	—	—	..	—	5
−·15	—	—	—	—	—	—	—	—	1	1	1	1	6	3	—	1	1	—	—	—	—	—	—	—	—	—	—	..	—	6
−·12	—	—	—	—	—	—	—	—	—	—	—	4	7	5	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	14
−·09	—	—	—	—	—	—	3	—	—	—	—	—	9	4	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	17
−·06	—	—	1	—	—	—	—	—	1	—	1	3	15	7	5	—	1	—	—	—	—	—	1	—	—	—	—	..	—	17
−·03	1	—	1	—	—	—	—	1	—	1	2	3	16	7	8	1	—	—	—	—	3	—	—	—	—	—	—	..	—	35
0	—	—	—	—	2	—	3	—	2	1	7	2	29	9	6	2	2	—	1	—	—	2	—	—	—	—	1	..	—	44
·03	—	—	—	—	1	—	—	—	—	—	—	8	8	10	—	3	1	1	—	1	—	—	—	—	—	1	—	..	—	69
·06	—	—	—	—	—	—	—	—	1	—	—	1	6	9	6	3	1	1	—	—	—	—	—	—	1	—	—	..	—	34
·09	—	—	—	—	—	—	—	—	—	—	1	3	5	7	3	3	—	—	—	—	—	—	—	—	—	1	—	..	—	29
·12	—	—	—	—	—	—	—	1	—	1	—	—	3	3	4	3	—	—	—	—	—	—	—	—	—	—	—	..	—	23
·15	—	—	—	—	—	—	—	—	—	—	2	1	2	2	—	2	—	—	—	—	—	—	—	—	—	—	—	..	—	15
·18	—	—	—	—	—	—	—	—	—	—	—	3	—	4	—	1	—	—	—	—	—	—	—	—	—	—	—	..	—	9
·21	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	1	—	—	—	—	—	—	—	—	—	—	—	..	—	8
·24	—	—	—	—	—	—	—	—	—	—	—	1	2	—	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	2
·27	—	—	—	—	—	—	—	—	—	—	—	—	3	2	—	—	—	1	—	—	—	—	—	—	—	—	1	..	—	4
·30	—	—	—	—	—	—	—	—	—	—	1	—	3	1	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	7
·33	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	1	—	—	—	—	1	—	—	..	—	6
·36	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	1	—	—	—	—	—	—	—	..	—	3
·39	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	2
·42	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	0
·45	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	1
·48	—	—	—	—	—	—	—	—	—	—	1	1	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	..	—	0
·51	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	3
·69	..	..	..	..	..	..	..	..	..	..	..	1	—	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	1
·81	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	2
·84	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	1
·87	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	0
·96	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	1	..	..	..	..	..	..	..	..	..	1
1·05	..	..	..	..	..	..	..	..	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	2
1·08	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	..	—	1
	1	1	4	0	3	0	8	2	6	6	19	34	128	80	44	23	10	5	4	2	4	2	1	1	3	1	2	..	1	395

LATERALITY OF VISUAL ACUITY & REFRACTION IN THE HORIZONTAL MERIDIAN

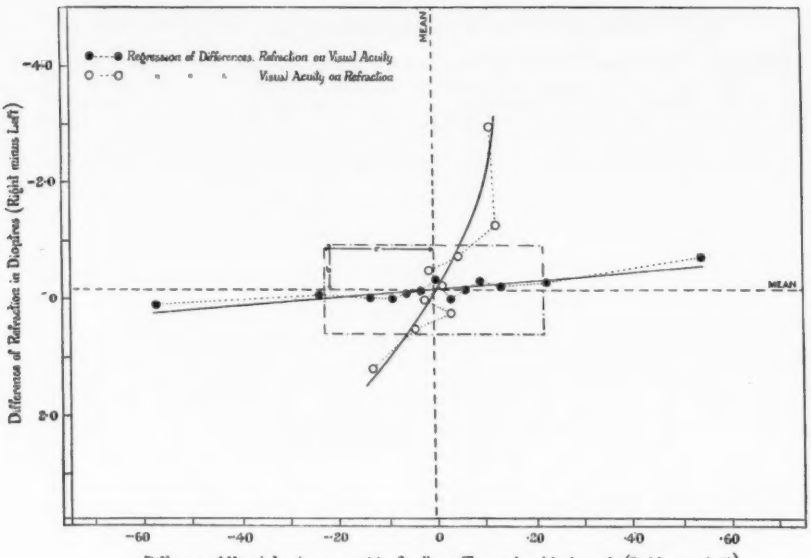


Diagram I.

LATERALITY OF VISUAL ACUITY & REFRACTION IN THE VERTICAL MERIDIAN

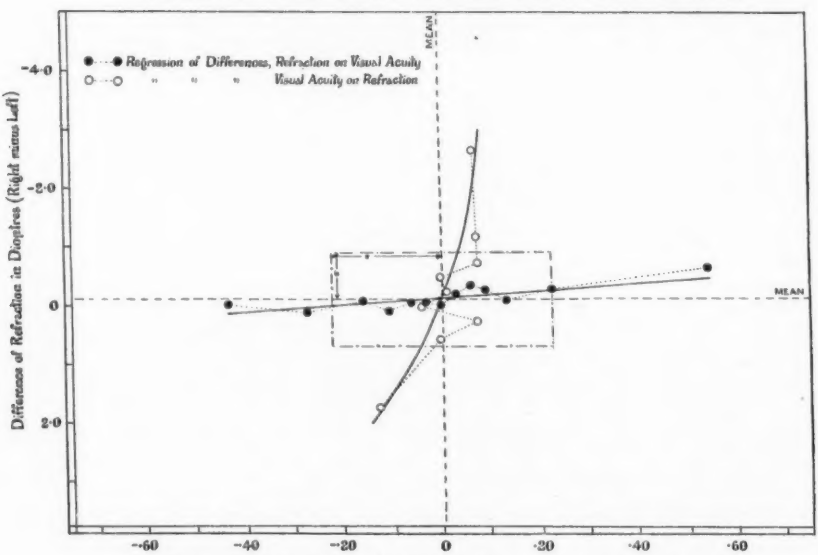


Diagram II.

Difference of Refraction in the Horizontal Meridian in Dioptres. (Corrected Values.) Right minus Left.

These values are all insignificant if compared with their probable errors, and accordingly we cannot assert that the eye dominant for Visual Acuity will be dominant for General Astigmatism.

(3) Correlation of the Differences of the two Refractions.

The data from which the following correlation values have been calculated are given in Table VII:

$r_{B_3B_4} = \cdot6642 \pm \cdot0184\ (383)$,

$\begin{cases} B_3 \text{ on } B_4: & \eta = \cdot6700, \quad \eta^2 = \cdot448941, \quad \overline{\eta}^2 = \cdot020888 \pm \cdot006961, \\ B_4 \text{ on } B_3: & \eta = \cdot6710, \quad \eta^2 = \cdot450256, \quad \overline{\eta}^2 = \cdot018277 \pm \cdot006521. \end{cases}$

TABLE VII.

Laterality of Refractions in the Horizontal and Vertical Meridians.

Difference of Refraction in the Vertical Meridian in Dioptres. (Central Values.) Right minus Left.

Difference of Refraction in the Horizontal Meridian in Dioptres. (Central Values.) Right minus Left.

	+2·75	+2·50	+2·25	+2·00	+1·75	+1·50	+1·25	+1·00	+0·75	+0·50	+0·25	0	−0·25	−0·50	−0·75	−1·00	−1·25	−1·50	−1·75	−2·00	−2·25	−2·50	−2·75	−3·00	−3·25	−3·50	..	−4·50	
+3·50	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	1
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
+2·75	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	1
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
+2·00	—	—	—	1	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	2
+1·75	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	1
+1·50	—	1	—	—	—	2	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	4
+1·25	—	1	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	2
+1·00	—	—	—	—	—	3	—	2	—	—	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	7
+0·75	—	—	—	—	—	—	—	—	2	2	2	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	8
+0·50	—	—	—	—	—	1	—	—	—	3	4	4	6	2	1	—	—	—	—	1	—	—	—	—	—	—	..	—	22
+0·25	—	—	—	—	—	—	1	—	3	—	12	12	6	1	4	—	—	—	—	—	—	—	1	—	—	—	..	—	40
0	—	—	—	—	—	1	—	1	—	8	4	76	9	8	2	—	—	—	—	—	—	—	—	1	—	—	..	—	110
−0·25	—	—	—	—	—	—	—	—	—	2	4	16	35	6	1	1	—	—	—	—	—	—	—	—	—	—	..	—	65
−0·50	—	—	—	—	—	—	1	1	—	1	4	13	12	22	3	2	—	—	—	—	—	—	—	—	—	—	..	—	59
−0·75	—	—	—	—	—	—	1	—	—	—	1	3	7	1	10	2	2	1	—	—	—	—	—	—	—	—	..	—	28
−1·00	—	—	—	—	—	—	—	—	—	—	—	—	1	1	—	3	1	—	—	—	—	—	—	—	—	—	..	—	6
−1·25	—	—	—	—	—	—	—	—	—	—	—	—	—	2	1	1	1	—	—	—	—	—	—	—	—	—	..	1	6
−1·50	—	—	—	1	—	—	—	—	—	1	—	—	1	—	—	—	—	1	1	1	—	—	—	—	—	—	..	—	6
−1·75	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	..	—	1
−2·00	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	1	—	—	1	—	—	—	—	..	—	3
−2·25	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	0
−2·50	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	..	—	1
−2·75	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	..	—	1
−3·00	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	2	—	1	..	—	5
−3·25	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2	—	—	..	—	2
−3·50	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	..	—	1
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
−4·25	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	..	—	1
−4·50	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	0
	1	3	0	2	0	8	3	6	7	17	32	127	78	43	22	10	4	3	2	2	2	1	1	5	1	2	..	1	383

All these values are significant. They show that an individual with a superior horizontal refraction of one eye will also show a superior vertical refraction of the same eye. In Diagram III the two regression lines are linear and have been fitted with the following equations:

(*a*) Probable $B_3 = \cdot64344B_4 + \cdot08975$.

(*b*) Probable $B_4 = \cdot68563B_3 + \cdot03513$.

(4) Difference of General Astigmatism with Differences of the two Refractions.

The last two pairs of ocular characters to be studied are General Astigmatism with each of the two Refractions. The following constants can be deduced from the tabulated data:

$r_{B_3B_5} = \cdot0843 \pm \cdot0337\,(394)$, $\begin{cases} B_3 \text{ on } B_5: & \eta^2 = \cdot027564, & \overline{\eta}^2 = \cdot015228 \pm \cdot005878, \\ B_5 \text{ on } B_3: & \eta^2 = \cdot025033, & \overline{\eta}^2 = \cdot017766 \pm \cdot006341, \end{cases}$

$r_{B_4B_5} = \cdot1039 \pm \cdot0336\,(395)$, $\begin{cases} B_4 \text{ on } B_5: & \eta^2 = \cdot017729, & \overline{\eta}^2 = \cdot015190 \pm \cdot005862, \\ B_5 \text{ on } B_4: & \eta^2 = \cdot051185, & \overline{\eta}^2 = \cdot020253 \pm \cdot006751. \end{cases}$

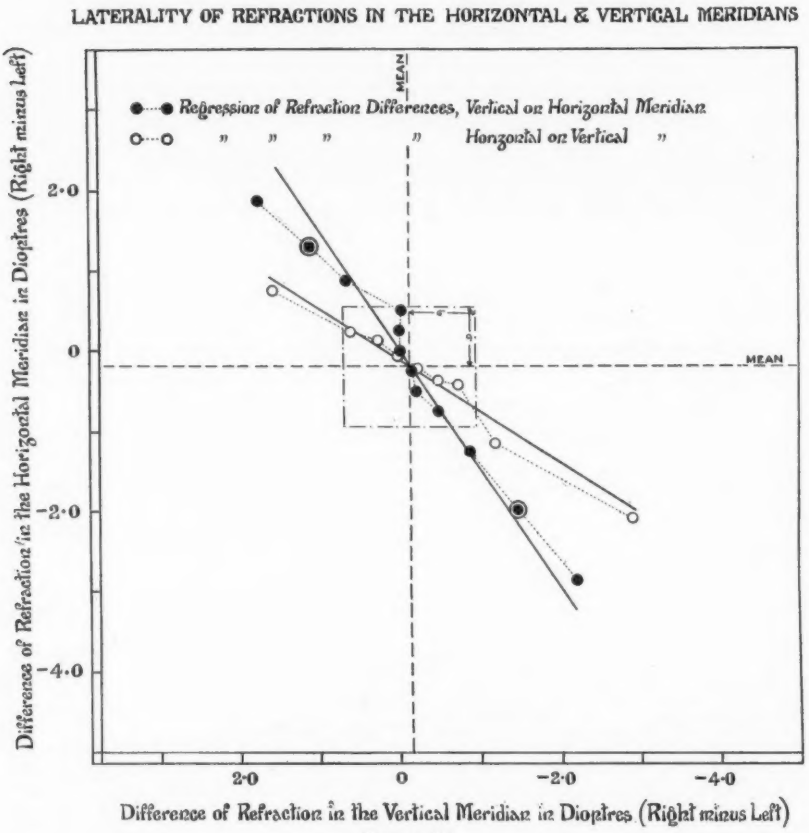


Diagram III.

It will be seen that the first $r_{B_3B_5}$ is only 2·2 times, and the second $r_{B_4B_5}$ is nearly three times its probable error thus showing probably a slight degree of significance. Only the last of the four correlation ratios is likely to be significant because here the value of $\eta^2_{B_5B_4}$ minus $\overline{\eta}^2_{B_5B_4}$ is about 4·5 times the probable error of $\overline{\eta}^2_{B_5B_4}$.

From the above results we may conclude that there is a high correlation between the dominance of the two Refractions and a moderate correlation between the dominance of Visual Acuity and that of each of the two Refractions. The other correlations do not really give any value of appreciable significance. Therefore it is clear from these results that it is not true that the lateralities of any two ocular characters are always highly associated.

(*c*) *On the Quantitative Interrelations of the Dominance of the Manual Characters.*

There are fifteen correlation coefficients for all the possible pairs of the six manual characters. For convenience we shall study them in groups.

(1) Difference of Grip (A_1) with those of the three Pulling Tests (A_2, A_3 and A_4), Steadiness (A_5) and Balancing (A_6).

The data for the comparison of Grip and First Pull are given in Table VIII. We can calculate the following correlation values:

$r_{A_1A_2} = \cdot1686 \pm \cdot0308$ (454),

A_1 on A_2: $\eta = \cdot2624$, $\eta^2 = \cdot068868$, $\bar{\eta}^2 = \cdot026432 \pm \cdot007173$,

A_2 on A_1: $\eta = \cdot2605$, $\eta^2 = \cdot067872$, $\bar{\eta}^2 = \cdot030837 \pm \cdot007730$.

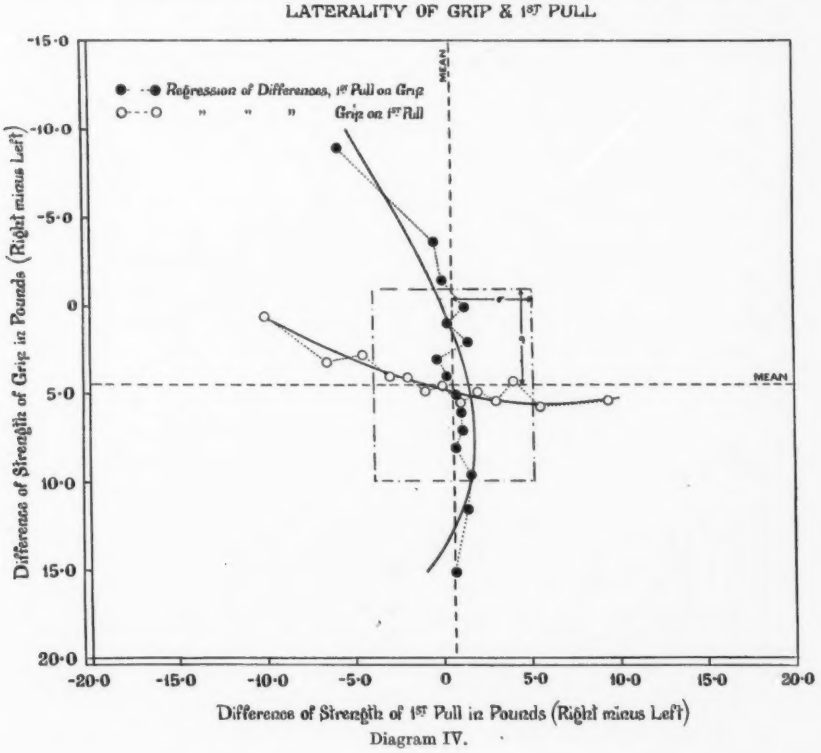


Diagram IV.

TABLE VIII.

Laterality of Strength of Grip and Strength of First Pull.

Difference of Strength of First Pull in Pounds. (Central Values.) Right minus Left.

	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0	−1	−2	−3	−4	−5	−6	−7	−8	−9	−10	−11	−12	−13	:	−17		
1																													1								1
3			1							1																										2
6										1		1																								2
9											2		1																							3
9												1	1	1																						4
14												1	1	2	1	1																				7
19													1	1	2	2	2	1																		9
31														1	1	3	4	3	1	1																15
41															1	2	3	5	4	4	2	1														23
47																1	3	2	3	5	4	2	1													25
50																	1	1	4	5	2	2	2	1												25
44																		1	2	1	4	5	2	3	3	2	2	4								22
32																			1	4	4	3	3	2	2	3	3	1								16
43																				2	1	4	1	2	3	2	1	2	6	2	4					21
23																					1	1	4	2	2	1	1	1		1							11
17																						1	2	1	1	1	2	1								8
15																							1	1	2	1	2	1	1							7
13																								1	1	2	1	1	2							6
12																									1	1	1	2	1							5
4																										1		1								2
7																											1	1	2							4
2																												1								1
0																																				0
2																													1							1
1																																				0
..																																				0
:																																				0
−17																																				1
	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0	−1	−2	−3	−4	−5	−6	−7	−8	−9	−10	−11	−12	−13	:	−17	454

Difference of Strength of Grip in Pounds. (Central Values.) Right minus Left.

The value of $r_{A_1A_2}$ is small but significant. The two correlation ratios are practically identical and are larger than $r_{A_1A_2}$; they are clearly significant. The two skew regression lines are shown in Diagram IV. They were fitted with cubic equations by the least square method and give:

(*a*) Probable $A_1 = 4{\cdot}70583 + {\cdot}263401A_2 - {\cdot}0180623A_2^2 - {\cdot}00039726A_2^3$.

(*b*) Probable $A_2 = -{\cdot}00083 + {\cdot}439087A_3 - {\cdot}0187809A_3^2 - {\cdot}00099520A_3^3$.

From the above diagram it will be seen that the cubics for the probable dominance of grip to pull and pull to grip are not similar. A maximum A_1 of 5·54610 lbs. is found when A_2 is 6·07424 lbs. while the maximum A_2 of 1·81488 lbs. is found when A_1 is 7·36939 lbs.

The correlation constants for difference of Grip with those of the remaining four manual characters are found to be:

$r_{A1A_3} = {\cdot}0842 \pm {\cdot}0316$ (450), $\begin{cases} A_1 \text{ on } A_3: & \eta^2 = {\cdot}051249, & \overline{\eta}^2 = {\cdot}028889 \pm {\cdot}007523, \\ A_3 \text{ on } A_1: & \eta^2 = {\cdot}038839, & \overline{\eta}^2 = {\cdot}031111 \pm {\cdot}007798, \end{cases}$

$r_{A1A_4} = {\cdot}0909 \pm {\cdot}0316$ (447), $\begin{cases} A_1 \text{ on } A_4: & \eta^2 = {\cdot}017833, & \overline{\eta}^2 = {\cdot}031320 \pm {\cdot}007851, \\ A_4 \text{ on } A_1: & \eta^2 = {\cdot}032020, & \overline{\eta}^2 = {\cdot}031320 \pm {\cdot}007851. \end{cases}$

$r_{A_1A_5} = {\cdot}0480 \pm {\cdot}0314$ (459), $\begin{cases} A_1 \text{ on } A_5: & \eta^2 = {\cdot}017382, & \overline{\eta}^2 = {\cdot}019608 \pm {\cdot}006167, \\ A_5 \text{ on } A_1: & \eta^2 = {\cdot}034848, & \overline{\eta}^2 = {\cdot}030501 \pm {\cdot}007648, \end{cases}$

$r_{A_1A_6} = {\cdot}0525 \pm {\cdot}0316$ (454), $\begin{cases} A_1 \text{ on } A_6: & \eta^2 = {\cdot}015760, & \overline{\eta}^2 = {\cdot}019824 \pm {\cdot}006233, \\ A_6 \text{ on } A_1: & \eta^2 = {\cdot}034665, & \overline{\eta}^2 = {\cdot}030837 \pm {\cdot}007730. \end{cases}$

The values of the correlation coefficients for the dominance of Grip with the Last Pull and Lack of Endurance of Pull are a little less than three times their probable errors but may possibly be significant; the last two coefficients are obviously insignificant. All the correlation ratios except the first one are too small to allow of any significant skew association between the dominance of Grip and the four manual characters here studied.

(2) Difference of First Pull (A_2) with those of Last Pull (A_3) and Lack of Endurance (A_4).

The First Pull test is designed to measure the initial muscular strength of the hand, while the Last Pull gives a value for the muscular strength after thirty seconds continuous pulling; the difference in pulling gives the muscular strength lost. Since the interrelations between the dominance of these three characters are important we will discuss them in detail.

The correlation constants are calculated from the data given in Tables IX, X and XI and set forth below:

$r_{A_2A_3} = {\cdot}2357 \pm {\cdot}0301$ (452),

$\begin{cases} A_2 \text{ on } A_3: & \eta = {\cdot}3198, & \eta^2 = {\cdot}102271, & \overline{\eta}^2 = {\cdot}028761 \pm {\cdot}007490, \\ A_3 \text{ on } A_2: & \eta = {\cdot}2752, & \eta^2 = {\cdot}075758, & \overline{\eta}^2 = {\cdot}026549 \pm {\cdot}007205, \end{cases}$

$r_{A_2A_4} = {\cdot}6058 \pm {\cdot}0201$ (450),

$\begin{cases} A_2 \text{ on } A_4: & \eta = {\cdot}6127, & \eta^2 = {\cdot}375398, & \overline{\eta}^2 = {\cdot}031111 \pm {\cdot}007798, \\ A_4 \text{ on } A_2: & \eta = {\cdot}6059, & \eta^2 = {\cdot}367056, & \overline{\eta}^2 = {\cdot}028889 \pm {\cdot}007523, \end{cases}$

$r_{A_3A_4} = -{\cdot}6357 \pm {\cdot}0189$ (450),

$\begin{cases} A_3 \text{ on } A_4: & \eta = {\cdot}6359, & \eta^2 = {\cdot}404314, & \overline{\eta}^2 = {\cdot}033333 \pm {\cdot}008062, \\ A_4 \text{ on } A_3: & \eta = {\cdot}6351, & \eta^2 = {\cdot}403408, & \overline{\eta}^2 = {\cdot}028889 \pm {\cdot}007523. \end{cases}$

TABLE IX.

Laterality of First Pull and Last Pull.

Difference of Strength of Last Pull in Pounds. (Central Values.) Right minus Left.

Difference of Strength of First Pull in Pounds. (Central Values.) Right minus Left.

	−20	..	−17	−16	−15	−14	−13	−12	−11	−10	−9	−8	−7	−6	−5	−4	−3	−2	−1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	..	19	
−17	—	..	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
..	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..
−13	—	..	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
−12	—	..	—	—	—	—	—	—	—	—	—	—	1	—	1	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	3
−11	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
−10	—	..	—	—	—	—	—	—	—	—	—	—	1	—	—	—	1	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2
−9	—	..	—	—	—	—	—	—	—	—	—	—	—	—	1	1	1	—	—	1	1	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	6
−8	—	..	—	—	—	—	1	—	—	—	—	—	—	1	2	1	—	1	—	—	—	1	—	1	—	—	—	—	—	—	—	—	—	—	—	—	3
−7	—	..	—	—	—	—	—	—	—	1	—	—	2	1	1	1	1	1	1	1	2	3	—	1	1	—	—	—	—	—	—	—	—	—	—	—	12
−6	—	..	—	—	—	—	—	—	—	1	—	—	1	—	1	1	1	2	—	1	2	—	1	—	1	—	—	—	1	—	—	—	—	—	—	—	—	13
−5	—	..	—	—	—	—	—	—	—	1	—	—	1	—	—	2	1	2	—	1	1	1	1	—	1	1	—	—	—	—	—	—	—	—	—	—	—	14
−4	—	..	—	—	1	—	—	—	—	—	—	—	—	1	1	3	1	1	1	2	2	2	1	1	3	3	—	—	—	—	—	—	—	—	—	—	18
−3	—	..	—	—	—	1	—	—	—	—	—	—	—	—	2	3	3	1	1	2	3	2	2	3	3	4	1	—	—	—	—	—	—	—	—	—	23
−2	—	..	—	—	1	—	—	1	—	—	—	—	—	—	4	5	3	1	2	3	5	5	3	5	4	2	1	—	—	—	—	—	—	—	—	—	43
−1	—	..	—	1	—	—	—	—	1	1	—	—	—	4	—	4	2	3	3	6	5	4	6	3	1	4	2	1	—	—	—	—	—	—	—	—	32
0	—	..	1	—	—	—	—	—	—	—	—	—	—	—	1	5	3	6	6	5	4	5	4	5	6	1	2	1	2	—	—	—	1	—	—	—	45
1	—	..	—	—	—	—	—	—	—	—	—	—	—	1	—	3	6	6	2	6	4	6	5	4	7	8	3	1	1	2	—	—	—	—	1	—	51
2	—	..	—	1	—	—	—	—	—	—	—	—	—	1	3	—	—	3	2	6	7	—	7	6	6	3	3	—	1	1	—	—	—	—	—	—	47
3	—	..	—	—	—	—	—	—	—	—	—	1	—	4	—	2	—	2	2	3	4	3	2	2	2	3	2	—	—	2	—	—	—	1	—	—	40
4	—	..	—	—	—	—	—	1	—	—	—	—	—	3	—	—	1	2	—	2	2	4	2	2	1	2	1	—	—	—	—	—	—	—	—	—	30
5	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	2	2	4	2	3	1	2	—	—	—	1	—	—	—	—	—	—	19
6	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	1	1	1	2	—	1	2	—	—	—	—	—	—	—	—	—	—	—	15
7	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	1	2	1	1	2	1	—	—	—	—	—	—	—	—	—	—	—	—	9
8	—	..	—	—	—	—	—	—	—	—	—	1	—	—	—	—	1	1	2	—	—	—	—	1	1	—	2	—	—	—	—	—	—	—	—	—	9
9	—	..	—	—	—	—	—	—	—	—	—	—	1	1	—	—	—	1	—	—	—	—	2	—	1	—	—	—	—	—	—	—	—	—	—	1	6
10	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	3
11	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
12	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2
13	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
14	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
15	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
16	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	1	—	—	—	—	—	—	—	—	—	—	—	2
17	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
18	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
19	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	
	1	..	1	1	2	1	1	2	1	4	1	5	9	17	22	30	22	38	40	51	37	46	35	33	23	9	6	6	1	2	1	0	1	2	..	1	452

TABLE X.

Laterality of First Pull and Lack of Endurance of Pull (First Pull – Last Pull).

Laterality of Lack of Endurance of Pull in Pounds. (Central Values.) Right minus Left.

Laterality of First Pull in Pounds. (Central Values.) Right minus Left.

	−17	−16	−15	−14	−13	−12	−11	−10	−9	−8	−7	−6	−5	−4	−3	−2	−1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
−17	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
−13	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
−12	—	—	—	—	—	—	—	—	—	—	1	—	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	3
−11	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
−10	—	—	—	—	—	—	—	—	1	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2
− 9	—	—	1	—	—	—	—	1	1	—	—	—	1	1	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	6
− 8	—	—	—	—	—	—	—	—	—	—	—	—	—	1	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	3
− 7	—	—	—	—	—	—	—	1	—	2	1	1	1	1	—	1	—	2	—	1	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	12
− 6	—	1	—	—	—	—	1	—	—	3	2	1	—	—	1	1	—	1	1	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	13
− 5	—	—	—	2	—	—	—	—	1	1	—	1	1	—	2	1	3	1	—	1	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	15
− 4	—	—	—	—	—	—	—	—	1	1	—	—	1	2	2	1	1	3	2	1	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	16
− 3	1	—	—	—	—	—	—	—	—	3	3	—	1	1	3	2	2	1	3	—	1	—	1	—	—	—	—	—	1	—	—	—	—	—	—	23
− 2	—	—	—	—	—	—	—	—	—	—	2	3	3	3	5	5	5	2	3	5	4	—	1	—	—	1	—	—	—	1	—	—	—	—	—	—	43
− 1	—	—	—	—	—	—	—	—	—	1	—	—	4	1	6	1	4	4	3	1	3	2	—	—	1	—	—	1	—	—	—	—	—	—	—	—	32
0	—	—	—	—	—	—	—	—	—	—	2	1	3	5	1	3	3	5	5	3	2	4	2	4	—	—	—	—	—	—	—	—	—	—	1	—	44
1	—	—	—	—	1	—	—	1	—	—	1	—	3	1	4	3	5	4	4	5	5	3	5	1	1	—	—	—	1	1	1	—	—	—	—	—	51
2	—	—	—	—	—	—	—	—	—	—	—	1	1	1	3	3	7	7	6	5	2	6	3	—	—	1	—	—	—	—	—	—	—	—	—	1	47
3	—	—	—	—	—	—	—	1	—	—	—	—	1	—	—	2	3	6	6	—	7	6	3	—	—	3	1	—	—	—	—	—	—	—	—	—	39
4	—	—	—	—	—	—	—	—	—	—	—	—	—	2	1	—	3	1	2	2	3	4	3	2	—	2	—	4	—	1	—	—	—	—	—	—	30
5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	3	2	4	1	2	—	2	1	—	—	3	—	—	—	—	—	1	—	20
6	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	2	—	3	2	4	2	1	—	—	—	—	—	—	—	—	—	—	—	15
7	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	2	1	1	1	1	—	—	—	—	—	—	—	—	—	9
8	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2	—	1	1	—	—	2	1	1	—	—	—	—	—	1	—	—	9
9	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	2	—	—	—	1	—	—	—	1	1	—	—	6
10	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	1	—	—	—	—	—	—	3
11	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
12	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	2
13	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
14	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	1
15	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	1
16	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
17	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	1	—	—	—	—	—	2
18	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
19	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	1
	1	1	1	2	1	0	1	5	5	11	12	8	23	20	31	25	36	37	40	28	35	33	29	11	8	9	4	7	8	6	3	3	1	2	2	1	450

TABLE XI.

Laterality of Last Pull and Lack of Endurance of Pull (First Pull – Last Pull).

Laterality of Lack of Endurance of Pull in Pounds (First Pull – Last Pull). (Central Values.) Right minus Left.

Laterality of Last Pull in Pounds. (Central Values.) Right minus Left	−17	−16	−15	−14	−13	−12	−11	−10	−9	−8	−7	−6	−5	−4	−3	−2	−1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
−20	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	1
−19	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	0
−18	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	1
−17	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
−16	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	—	—	—	—	—	0
−15	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2
−14	—	—	—	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	1	—	—	—	—	—	—	—	—	1
−13	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	2
−12	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	1
−11	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	4
−10	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	—	—	1	—	—	1	—	—	—	—	—	—	1
−9	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	1	—	—	1	—	—	—	—	—	—	1	—	—	1	—	—	—	5
−8	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	2	1	1	—	—	—	1	—	1	—	—	—	—	1	—	—	1	—	—	9
−7	—	—	—	—	—	—	—	1	—	—	—	—	1	—	1	—	—	1	2	1	—	—	1	—	4	—	1	1	4	3	—	—	—	—	1	—	—	17
−6	—	—	—	—	—	—	—	—	—	—	1	—	—	1	2	1	—	1	2	—	1	4	2	3	1	—	3	1	—	—	—	—	—	—	—	—	—	21
−5	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	1	3	3	3	3	3	4	4	4	—	—	2	—	—	—	—	—	—	—	—	—	30
−4	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	1	1	1	2	3	5	3	4	—	—	—	1	—	1	1	—	—	—	—	—	—	22
−3	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	1	2	2	3	5	1	3	3	3	2	—	—	—	1	1	1	1	—	—	—	—	37
−2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2	2	1	2	4	5	3	5	5	6	2	1	1	—	2	—	—	—	—	—	—	—	40
−1	—	—	—	—	—	—	—	—	1	—	—	1	—	—	2	2	2	5	4	5	4	5	5	6	7	4	1	—	—	—	1	—	1	—	—	—	50
0	—	—	—	—	—	—	—	—	1	—	1	1	1	2	3	5	4	5	4	5	5	7	4	1	2	1	—	—	—	—	1	—	—	1	—	—	38
1	—	—	—	—	—	—	—	1	—	2	2	1	1	1	5	1	3	5	4	6	—	3	1	4	1	—	—	—	—	—	—	—	—	—	—	—	46
2	—	—	—	—	—	—	—	—	—	3	—	—	1	3	6	3	5	6	7	2	2	4	2	2	—	2	—	—	—	—	—	—	1	—	—	—	35
3	—	—	—	—	—	—	—	1	—	1	—	—	3	1	1	4	3	7	2	2	3	1	1	1	—	—	—	—	—	—	—	—	—	—	—	—	33
4	—	—	—	—	—	—	—	—	1	1	3	3	4	5	1	4	3	3	1	2	—	1	1	—	—	—	—	—	—	—	1	—	—	—	—	—	23
5	—	—	—	—	—	—	1	—	1	3	2	—	3	1	3	3	3	—	—	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	9
6	—	—	1	—	—	—	—	—	—	—	—	1	3	1	—	—	—	—	—	—	2	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	5
7	—	—	—	—	—	—	—	—	—	1	2	—	1	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	6
8	—	—	—	—	—	—	—	—	—	—	2	1	1	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2
9	—	—	—	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2
10	—	1	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
11	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
12	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
13	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2
14	1	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	1
19	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	
	1	1	1	2	1	0	1	5	5	11	13	8	23	20	31	25	36	37	40	28	35	33	29	11	7	9	4	7	8	6	2	3	1	3	2	1	450

It will be seen that all the coefficients are large and significant, indicating a distinct association between the dominance of initial and final muscular strengths and of the lack of endurance of strength in pulling. The first coefficient shows that a dominance of one hand for First Pull is associated to a moderate extent with the dominance for Last Pull while the second coefficient tells us that the greater the dominance of the initial muscular strength the greater the dominance of final strength will be decreased. $r_{A_2A_4}$ with a negative sign indicates, what we might expect, that the greater the dominance of Last Pull the less will be the dominance of Lack of Endurance in Pulling.

The three pairs of correlation ratios are all significant and the regression lines are shown in Diagrams V, VI and VII. The regression lines of A_3 on A_2 and A_4 on A_2 shown in Diagrams V and VI are slightly skew and therefore will be fitted

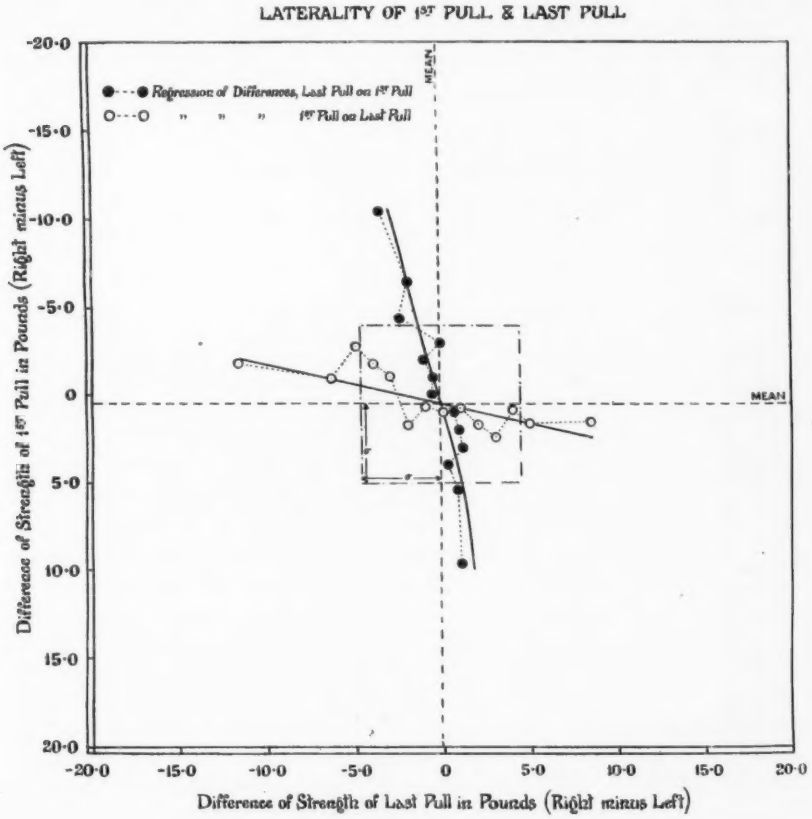


Diagram V.

with cubics while the others being linear have been fitted with linear regression equations. These six equations are given below:

(*a*) Probable $A_2 = \cdot 22387 A_3 + \cdot 48837$ (in terms of A_3).

(*b*) Probable $A_2 = \cdot 49872 A_4 + \cdot 17982$ („ „ „ A_4).

(*c*) Probable $A_3 = -\cdot 30691 + \cdot 286351 A_2 - \cdot 0011399 A_2^2 - \cdot 00064529 A_2^3$ („ „ „ A_2).

(*d*) Probable $A_3 = -\cdot 53037 A_4 + \cdot 20147$ („ „ „ A_4).

(*e*) Probable $A_4 = +\cdot 05952 + \cdot 726313 A_2 + \cdot 0135407 A_2^2 - \cdot 00002224 A_2^3$ („ „ „ A_2).

(*f*) Probable $A_4 = -\cdot 76194 A_3 + \cdot 57460$ („ „ „ A_3).

In order to throw more light on the nature and extent of muscular endurance in pulling, it may be worth while to study the correlations of the *absolute*

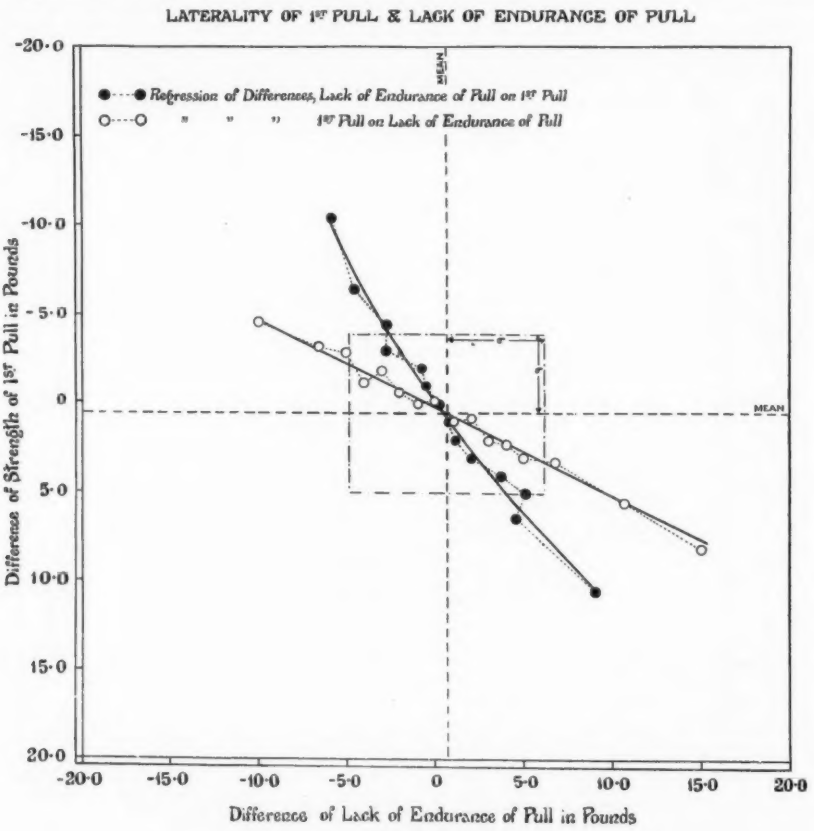


Diagram VI.

measures of the two given variates for right and left hand respectively. Tables XII and XIII (p. 113) provide the data for both variates upon which the following numerical results are based, 2 lbs. being taken as a working unit for each variate.

Characters	Mean	S.D.
Strength of First Pull for right hand (a_{2R}) 453 cases	40·096 ± ·229 lbs.	7·221 ± ·162 lbs.
Lack of Endurance in pulling for right hand (a_{4R}) 453 cases	15·089 ± ·173 lbs.	5·462 ± ·122 lbs.
Strength of First Pull for left hand (a_{2L}) 458 cases	39·587 ± ·207 lbs.	7·204 ± ·161 lbs.
Lack of Endurance in pulling for left hand (a_{4L}) 458 cases	14·369 ± ·166 lbs.	5·259 ± ·117 lbs.

$r_{a_{2R}, a_{4R}} = ·5639 \pm ·0216$, $r_{a_{2L}, a_{4L}} = ·4279 \pm ·0257$

(a_{2R} on a_{4R})

$\eta = ·5829$, $\eta^2 = ·339803$, $\bar{\eta}^2 = ·022075 \pm ·006800$

(a_{4R} on a_{2R})

$\eta = ·5727$, $\eta^2 = ·327960$, $\bar{\eta}^2 = ·024283 \pm ·006892$

(a_{2L} on a_{4L})

$\eta = ·4598$, $\eta^2 = ·211451$, $\bar{\eta}^2 = ·019651 \pm ·006179$

(a_{4L} on a_{2L})

$\eta = ·4460$, $\eta^2 = ·198943$, $\bar{\eta}^2 = ·026201 \pm ·007111$

From these results it may be seen that the average strength of First Pull for the right hand is about ·5 lb. stronger than that for the left, and the reducing amount after exertion, i.e. Lack of Endurance in pulling, for the right hand is nearly ·7 lb. more than that for the left. The value of $100\bar{a}_4/\bar{a}_2$ ($\bar{a}_4$ = mean of Lack of Endurance and $\bar{a}_2$ = mean of First Pull) for the right hand is 37·6 per cent. and that for the left 36·3 per cent., which are nearly identical.

The values of the correlation coefficients in the case of both hands are considerable although they are less than the values of the correlations of their relative dominances. These high values indicate that the increase of the value of one variate gives a proportional increase to the other, no matter which hand is tested. The reason may be possibly accounted for thus: Subjects who exert excessive muscular power at first pull become exhausted in strength earlier and will be more easily affected by a factor of physical fatigue than those who give a lower record at first pulling. The two pairs of correlation ratios given here are also essentially large. Diagram VIII (p. 114) shows the changes of the two pairs of regression lines for each hand. The reader will observe in the Diagrams that the regressions of a_4 (Lack of Endurance) on a_2 (First Pull) for both hands follow more or less similar straight lines. On the other hand, the regression curve of a_2 on a_4 for the right hand

changes fairly steadily; while the amount of a_4 gradually increases through a range of 26 lbs. there is a nearly proportional increase of a_2 which totals

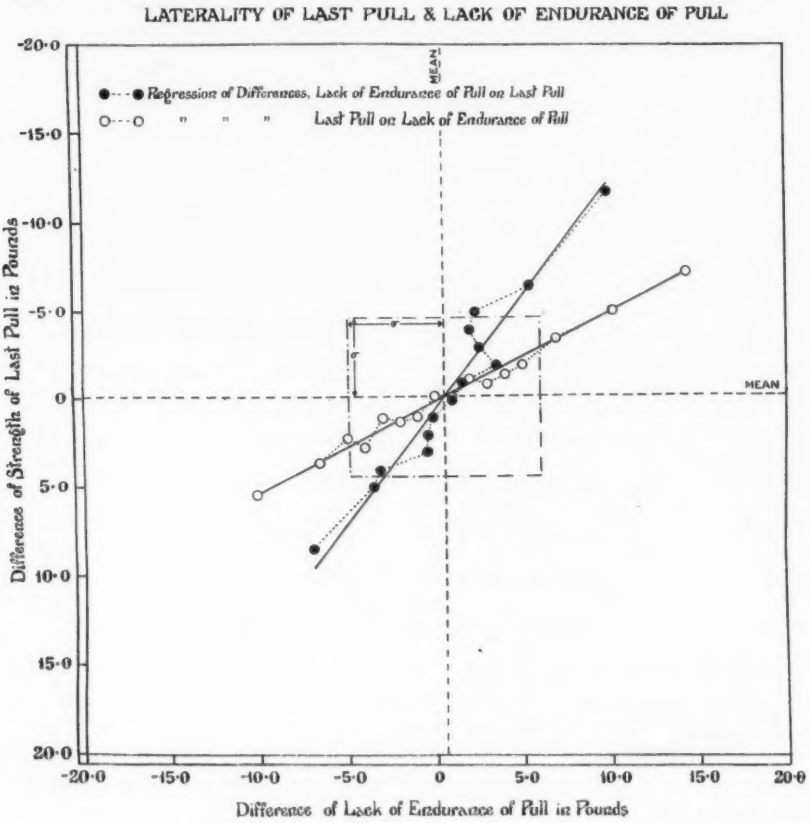


Diagram VII.

about 18 lbs. On the left, however, as the value of a_4 increases from 4 lbs. to 11·5 lbs., there corresponds a constant value of nearly 36 lbs. in a_2. After a_4 has increased up to 15·5 lbs., then the two variates correspond almost proportionally with one another in change.

According to the nature of their curvatures, the regression lines of a_2 on a_4 for both hands have been fitted with cubic and those of a_4 on a_2 with linear equations. The equations are given below:

(*a*) Probable $a_{2R} = 33{\cdot}19652 + {\cdot}202180a_{4R} + {\cdot}0123439a^2_{4R} + {\cdot}00012010a^3_{4R}$.

(*b*) Probable $a_{4R} = {\cdot}42669a_{2R} - 2{\cdot}02156$.

(*c*) Probable $a_{2L} = 49{\cdot}86460 - 3{\cdot}693951a_{4L} + {\cdot}3012458a^2_{4L} - {\cdot}00668708a^3_{4L}$.

(*d*) Probable $a_{4L} = {\cdot}31237a_{2L} + 2{\cdot}00321$.

TABLE XII.

First Pull and Difference of First Pull minus Last Pull in Pounds for Right Hand.

Strength of First Pull in Pounds for Right Hand. (Central Values.)

Difference of First Pull minus Last Pull in Pounds for Right Hand. (Central Values.)

Central Value	23·5	25·5	27·5	29·5	31·5	33·5	35·5	37·5	39·5	41·5	43·5	45·5	47·5	49·5	51·5	53·5	55·5	57·5	59·5	61·5	63·5	65·5	67·5	
−2·5	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
−0·5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
1·5	—	—	—	—	—	—	—	1	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	2
3·5	1	—	—	—	1	—	—	—	1	—	—	—	—	—	1	—	—	—	—	—	—	—	—	4
5·5	—	—	2	1	3	1	2	1	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	11
7·5	1	—	—	3	5	3	3	1	2	1	1	2	—	—	—	—	—	—	—	—	—	—	—	22
9·5	—	—	2	2	7	10	5	11	2	4	2	1	—	—	1	—	—	—	—	—	—	—	—	47
11·5	1	—	2	1	8	10	6	12	8	2	5	—	5	—	—	—	—	1	—	—	—	—	—	61
13·5	—	—	—	4	5	11	10	12	13	11	12	2	2	1	—	1	—	—	—	—	—	—	—	84
15·5	—	—	1	—	4	4	5	7	7	5	9	4	3	1	1	1	1	—	—	—	1	—	—	54
17·5	—	—	1	2	—	4	10	3	5	7	5	7	3	1	5	3	—	—	—	—	—	—	—	56
19·5	—	1	—	—	1	1	5	5	2	2	4	5	6	1	1	1	—	1	—	—	—	—	—	36
21·5	—	—	—	—	—	—	3	1	5	5	11	4	1	3	1	1	—	—	—	—	—	—	—	35
23·5	—	—	—	—	—	—	1	—	—	3	4	1	—	3	1	1	3	1	—	—	—	—	—	18
25·5	—	—	—	—	1	—	—	—	—	1	1	—	1	2	1	—	1	—	1	—	1	—	—	10
27·5	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	1	—	2	—	—	—	—	4
29·5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	1	—	—	1	3
31·5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	1	—	—	—	—	—	—	2
33·5	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	1	—	—	—	—	—	—	2
35·5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	1
	3	1	8	13	35	45	50	54	45	41	55	28	22	12	13	8	8	4	3	1	3	0	1	453

TABLE XIII.

First Pull and Difference of First Pull minus Last Pull in Pounds for Left Hand.

Strength of First Pull in Pounds for Left Hand. (Central Values.)

Difference of First Pull minus Last Pull in Pounds for Left Hand. (Central Values.)

Central Value	23·5	25·5	27·5	29·5	31·5	33·5	35·5	37·5	39·5	41·5	43·5	45·5	47·5	49·5	51·5	53·5	55·5	57·5	59·5	61·5	63·5	65·5	67·5	
−2·5	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	1
−0·5	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
1·5	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
3·5	—	—	—	1	—	1	—	1	1	1	—	—	1	—	—	—	—	—	—	—	—	—	—	6
5·5	—	1	1	—	4	1	1	—	2	1	1	—	1	—	—	—	—	—	—	—	—	—	—	13
7·5	—	1	3	7	3	4	3	5	—	1	4	2	—	1	—	1	—	—	—	—	—	—	—	35
9·5	—	2	1	2	4	6	13	3	1	3	2	3	2	—	—	—	—	—	—	—	—	—	—	42
11·5	1	1	4	7	7	9	15	4	9	9	7	2	1	—	—	—	—	—	—	—	—	—	—	76
13·5	—	—	1	—	5	4	10	12	8	7	9	3	1	—	3	—	1	—	—	1	—	—	—	65
15·5	1	1	1	2	3	7	9	12	7	7	10	5	3	2	3	1	—	—	—	—	1	—	—	75
17·5	—	—	—	1	4	2	7	7	5	8	4	7	4	1	1	3	4	—	1	—	1	—	—	60
19·5	—	—	—	—	1	1	2	2	2	5	5	3	3	—	2	—	—	—	—	—	—	—	1	27
21·5	—	—	—	—	—	2	3	1	3	1	3	2	4	5	2	—	1	—	—	—	—	—	—	27
23·5	—	—	—	—	—	2	1	1	—	—	2	1	2	1	—	—	3	—	—	—	—	—	—	13
25·5	—	—	—	—	—	—	—	—	2	1	1	1	1	—	1	—	—	—	—	—	—	—	—	7
27·5	—	—	—	—	—	—	—	—	—	—	—	1	—	—	2	—	—	—	—	—	—	—	—	3
29·5	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	1	—	—	—	—	—	—	2
31·5	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	1	—	2
33·5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	1
35·5	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	1
	2	6	12	19	31	39	65	48	41	44	49	32	23	11	14	5	11	0	1	1	2	1	1	458

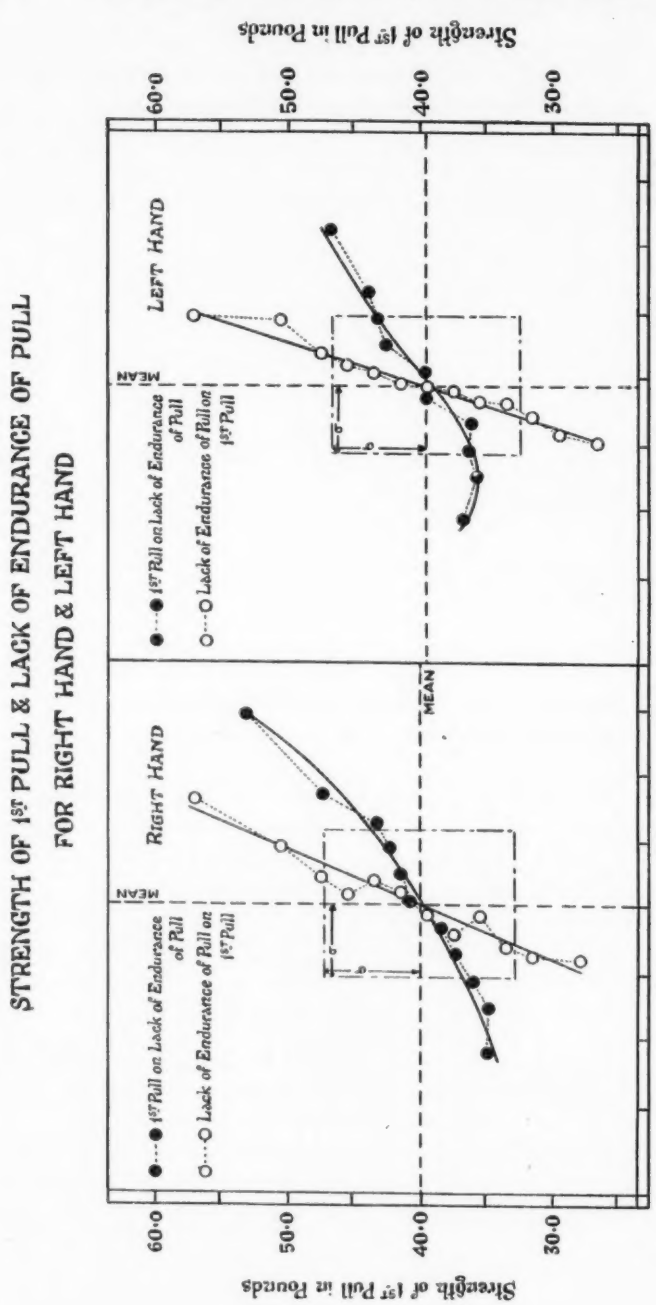


Diagram VIII.

(3) Differences of Three Pulling Tests (A_2, A_3 and A_4) with those of Steadiness of Hand (A_5) and Balancing (A_6).

Here we have to deal with the three differences of the three pulling tests for the two hands associated respectively with the differences for steadiness and balancing measurements for the same hands.

As usual all associational constants were obtained from the separate data. Their results are given below:

Product Moment Correlations	Correlation Ratios
$r_{A_2A_5} = \cdot0663 \pm \cdot0315$ (456),	A_2 on A_5: $\eta^2 = \cdot015549$, $\overline{\eta}^2 = \cdot019737 \pm \cdot006206$; A_5 on A_2: $\eta^2 = \cdot043202$, $\overline{\eta}^2 = \cdot028509 \pm \cdot007425$,
$r_{A_2A_6} = \cdot0695 \pm \cdot0316$ (451),	A_2 on A_6: $\eta^2 = \cdot011502$, $\overline{\eta}^2 = \cdot019956 \pm \cdot006275$; A_6 on A_2: $\eta^2 = \cdot023020$, $\overline{\eta}^2 = \cdot028825 \pm \cdot007508$,
$r_{A_3A_5} = \cdot0679 \pm \cdot0316$ (451),	A_3 on A_5: $\eta^2 = \cdot013987$, $\overline{\eta}^2 = \cdot019956 \pm \cdot006275$; A_5 on A_3: $\eta^2 = \cdot034959$, $\overline{\eta}^2 = \cdot028825 \pm \cdot007508$,
$r_{A_3A_6} = \cdot0045 \pm \cdot0320$ (444),	A_3 on A_6: $\eta^2 = \cdot020033$, $\overline{\eta}^2 = \cdot020270 \pm \cdot006371$; A_6 on A_3: $\eta^2 = \cdot048800$, $\overline{\eta}^2 = \cdot029279 \pm \cdot007622$,
$r_{A_4A_5} = \cdot0136 \pm \cdot0318$ (450),	A_4 on A_5: $\eta^2 = \cdot016597$, $\overline{\eta}^2 = \cdot020000 \pm \cdot006288$; A_5 on A_4: $\eta^2 = \cdot051382$, $\overline{\eta}^2 = \cdot033333 \pm \cdot008062$,
$r_{A_4A_6} = \cdot0628 \pm \cdot0319$ (444),	A_4 on A_6: $\eta^2 = \cdot012879$, $\overline{\eta}^2 = \cdot020270 \pm \cdot006371$. A_6 on A_4: $\eta^2 = \cdot035844$, $\overline{\eta}^2 = \cdot033784 \pm \cdot008169$.

All the correlation coefficients are insignificant. Again, none of the correlation ratios can be regarded as of definitely significant value, when compared with the corresponding value of $\overline{\eta}^2$, the correlation ratio for no association. Thus we are unable to assert that any important relationship exists between the dominances of any pair of these given six variates.

(4) Difference of Steadiness of Hand and Balancing (A_5 and A_6).

Lastly, let us turn to the dominance of Steadiness of Hand and Balancing. We deduce from the data for the given two characters the following values:

Product Moment Correlation $r_{A_5A_6} = \cdot0040 \pm \cdot0315$ (457).

Correlation Ratio of A_5 on A_6: $\eta^2 = \cdot036503$, $\overline{\eta}^2 = \cdot019694 \pm \cdot006194$,

" " A_6 on A_5: $\eta^2 = \cdot018632$, $\overline{\eta}^2 = \cdot019694 \pm \cdot006194$.

Thus, the correlation coefficient is much less than its probable error and, consequently, of no importance. This seems to indicate that any orderly change of dominance in Steadiness of Hand has little influence on that of Balancing. The correlation ratio of A_5 on A_6 is comparatively large but still hardly significant when the difference of $\eta^2_{A_5A_6}$ and $\overline{\eta}^2_{A_5A_6}$ is considered, i.e. only 2·7 times the probable error.

The foregoing results show that while for certain of the manual characteristics there is a considerable correlation between the dominant lateralities, yet for others there is no sensible relation at all. In other words we cannot assert that if the one hand is dominant over the other for one character, it will be to the

same extent dominant for all the other characters. There appears to be no such attribute of universal laterality of the hand. For some characters there is a fairly high correlation, for others none at all. An individual may have any degree of right- or left-handedness, and it would appear that he may be right-handed for one character and left-handed for a second.

It will be clearly seen from Table IV *b* that among 15 pairs of manual characters the value of opposite laterality varies from 21·11 per cent. to 60·88 per cent. giving a mean of 34·52 per cent. That is to say, there is at least one-third of all individuals who are dextral in one manual character but sinistral in the second. Even with such a high correlated pair of manual characters as First Pull and Lack of Endurance in pulling, there is still more than one-fifth of all persons who fail to obey the rule of absolute Manual Laterality.

Problem 4. *Is there any Truth in the Theory of Universal Laterality, i.e. is there any association between dominance in hand and eye?*

(*a*) *On the Relationships in Percentage Occurrences between the Dominance of Ocular and Manual Characters.*

Before we proceed to test the numerical relations between any pair of ocular and manual characters by the usual correlation methods, we will first study the correspondence of percentages between the observed and expected values in each category of laterality. The process followed is exactly the same as in the former cases. Table XV (pp. 115—125) provides the thirty combined distributions for laterality of any two variates of hand and eye. Three kinds of percentages for each type of laterality are given in the table. First, those of each cell on the total of the table, then of the values to be expected on the assumption of no association and lastly, the percentage of the cell frequencies on the column total. We note that:

(1) The differences between the percentages actually observed and those to be expected are practically very small. In most cases they only differ in the decimal places, which might easily occur owing to random sampling. The largest difference is merely 3·53 per cent. (observed = 42·30 per cent. and expected = 38·77 per cent.) in the case of manual dextrality and ocular ambilaterality in the Steadiness test associated with the Near Point. It is only about 1·52 times the probable error of their difference, and therefore not of significance. It may be remarked, however, that this is one of the cases in which the eye is auxiliary to the hand.

(2) When similar comparisons are made between the percentage in any column and the corresponding percentage in the marginal column, we find in 98 cases out of 100 the differences between them are insignificant. Those differences which are less than 10·0 per cent. in amount seem to indicate no significance. For instance, in the case of First Pull and Vertical Refraction (see Table XV, No. 9) in the category of manual dextrality and visual sinistrality, the value is 9·77 per cent. (63·10—52·33 per cent.) which is just equal to 2·41 times the probable error of

TABLE XIV.

Correlation Constants for (a) Lateralities of Ocular Characters, (b) Lateralities of Manual Characters.

(*a*) Lateralities of Ocular Characters.

Ocular Characters	Product Moment inter-Correlation Coefficients ($r_{B_xB_y}$)	Correlation Ratios of B_x on B_y		Correlation Ratios of B_y on B_x	
		Values of $\eta^2_{B_x \cdot B_y}$	Values of $\overline{\eta}^2_{B_x \cdot B_y}$ and their Probable Errors	Values of $\eta^2_{B_y \cdot B_x}$	Values of $\overline{\eta}^2_{B_y \cdot B_x}$ and their Probable Errors
Near Point and Visual Acuity	$r_{B_1B_2} = +\cdot1122 \pm \cdot0327$	·039241	·028986 ± ·007856	·023034	·012077 ± ·005115
Near Point and "Horizontal" Refraction ...	$r_{B_1B_3} = +\cdot1028 \pm \cdot0339$	·030875	·018041 ± ·006437	·029833	·012887 ± ·005455
Near Point and "Vertical" Refraction	$r_{B_1B_4} = +\cdot0069 \pm \cdot0342$	·034814	·020566 ± ·006854	·010406	·012853 ± ·005440
Near Point and General Astigmatism	$r_{B_1B_5} = +\cdot0206 \pm \cdot0333$	·029790	·014599 ± ·005636	·017959	·012165 ± ·005152
Visual Acuity and "Horizontal" Refraction ...	$r_{B_2B_3} = +\cdot2056 \pm \cdot0326$	·061832 ($\eta = \cdot2487$)	·020356 ± ·006786	·054325 ($\eta = \cdot2331$)	·030534 ± ·008268
Visual Acuity and "Vertical" Refraction ...	$r_{B_2B_4} = +\cdot2062 \pm \cdot0325$	·057399 ($\eta = \cdot2396$)	·020253 ± ·006751	·055094 ($\eta = \cdot2347$)	·030380 ± ·008226
Visual Acuity and General Astigmatism	$r_{B_2B_5} = +\cdot0612 \pm \cdot0330$	·018334	·014388 ± ·005556	·008572	·028777 ± ·007799
"Horizontal" Refraction and "Vertical" Refraction	$r_{B_3B_4} = +\cdot6642 \pm \cdot0184$	·448941 ($\eta = \cdot6700$)	·020888 ± ·006961	·450256 ($\eta = \cdot6710$)	·018277 ± ·006521
"Horizontal" Refraction and General Astigmatism	$r_{B_3B_5} = +\cdot0843 \pm \cdot0337$	·027564	·015228 ± ·005878	·025033	·017766 ± ·006341
"Vertical" Refraction and General Astigmatism ...	$r_{B_4B_5} = +\cdot1039 \pm \cdot0336$	·017729	·015190 ± ·005862	·051185	·020253 ± ·006751

(*b*) Lateralities of Manual Characters.

Manual Characters	Product Moment inter-Correlation Coefficients ($r_{A_xA_y}$)	Correlation Ratios of A_x on A_y		Correlation Ratios of A_y on A_x	
		Values of $\eta^2_{A_x \cdot A_y}$	Values of $\overline{\eta}^2_{A_x \cdot A_y}$ and their Probable Errors	Values of $\eta^2_{A_y \cdot A_x}$	Values of $\overline{\eta}^2_{A_y \cdot A_x}$ and their Probable Errors
Grip and First Pull	$r_{A_1A_2} = \cdot1686 \pm \cdot0308$	·068868 ($\eta = \cdot2624$)	·026432 ± ·007173	·067872 ($\eta = \cdot2605$)	·030837 ± ·007730
Grip and Last Pull	$r_{A_1A_3} = \cdot0842 \pm \cdot0316$	·051249	·028889 ± ·007523	·038839	·031111 ± ·007798
Grip and Lack of Endurance of Pull	$r_{A_1A_4} = \cdot0909 \pm \cdot0316$	·017833	·031320 ± ·007851	·032020	·031320 ± ·007851
Grip and Steadiness Test	$r_{A_1A_5} = \cdot0480 \pm \cdot0314$	·017382	·019608 ± ·006167	·034848	·030501 ± ·007648
Grip and Balancing Test	$r_{A_1A_6} = \cdot0525 \pm \cdot0316$	·015760	·019824 ± ·006233	·034665	·030837 ± ·007730
First Pull and Last Pull	$r_{A_2A_3} = \cdot2357 \pm \cdot0301$	·102271 ($\eta = \cdot3198$)	·028761 ± ·007490	·075758 ($\eta = \cdot2752$)	·026549 ± ·007205
First Pull and Lack of Endurance of Pull ...	$r_{A_2A_4} = \cdot6058 \pm \cdot0201$	·375398 ($\eta = \cdot6127$)	·031111 ± ·007798	·367056 ($\eta = \cdot6059$)	·028889 ± ·007523
First Pull and Steadiness Test	$r_{A_2A_5} = \cdot0663 \pm \cdot0315$	·015549	·019737 ± ·006206	·043202	·028509 ± ·007425
First Pull and Balancing Test	$r_{A_2A_6} = \cdot0695 \pm \cdot0316$	·011502	·019956 ± ·006275	·023020	·028825 ± ·007508
Last Pull and Lack of Endurance of Pull ...	$r_{A_3A_4} = -\cdot6357 \pm \cdot0189$	·404314 ($\eta = \cdot6359$)	·033333 ± ·008062	·403408 ($\eta = \cdot6351$)	·028889 ± ·007523
Last Pull and Steadiness Test	$r_{A_3A_5} = \cdot0679 \pm \cdot0316$	·013987	·019956 ± ·006275	·034959	·028825 ± ·007508
Last Pull and Balancing Test	$r_{A_3A_6} = \cdot0045 \pm \cdot0320$	·020033	·020270 ± ·006371	·048800	·029279 ± ·007622
Lack of Endurance of Pull and Steadiness Test ...	$r_{A_4A_5} = \cdot0136 \pm \cdot0318$	·016597	·020000 ± ·006288	·051382	·033333 ± ·008062
Lack of Endurance of Pull and Balancing Test ...	$r_{A_4A_6} = \cdot0628 \pm \cdot0319$	·012879	·020270 ± ·006371	·035844	·033784 ± ·008169
Steadiness Test and Balancing Test	$r_{A_5A_6} = \cdot0040 \pm \cdot0315$	·036503	·019694 ± ·006194	·018632	·019694 ± ·006194

TABLE XV.

Totals with Percentages observed and Percentages expected on the basis of no Association for each Category of Manual and Ocular Laterality.

(1) Grip and Near Point.

Manual Laterality for Grip	Ocular Laterality for Near Point									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	13 (3·16)	(2·31)	[19·70]	35 (8·52)	(8·56)	[14·29]	11 (2·68)	(3·49)	[11·00]	59 (14·36)
Ambilaterality	3 (0·73)	(0·98)	[4·55]	16 (3·89)	(3·62)	[6·53]	6 (1·46)	(1·48)	[6·00]	25 (6·08)
Dextrality	50 (12·17)	(12·77)	[75·75]	194 (47·20)	(47·43)	[79·18]	83 (20·19)	(19·36)	[83·00]	327 (79·56)
Totals	66 (16·06)			245 (59·61)			100 (24·33)			411 (100)

(2) Grip and Visual Acuity.

Manual Laterality for Grip	Ocular Laterality for Visual Acuity									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	26 (6·21)	(6·46)	[13·98]	11 (2·63)	(2·40)	[15·94]	24 (5·72)	(5·70)	[14·63]	61 (14·56)
Ambilaterality	13 (3·10)	(2·44)	[6·99]	3 (0·71)	(0·90)	[4·35]	7 (1·68)	(2·15)	[4·27]	23 (5·49)
Dextrality	147 (35·08)	(35·49)	[79·03]	55 (13·13)	(13·17)	[79·71]	133 (31·74)	(31·29)	[81·10]	335 (79·95)
Totals	186 (44·39)			69 (16·47)			164 (39·14)			419 (100

(3) Grip and Refraction in "Horizontal" Meridian.

Manual Laterality for Grip	Ocular Laterality for Refraction in "Horizontal" Meridian									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	15 (3·84)	(3·30)	[16·67]	16 (4·09)	(4·40)	[13·33]	25 (6·39)	(6·62)	[13·81]	56 (14·32)
Ambilaterality	6 (1·54)	(1·41)	[6·67]	9 (2·30)	(1·88)	[7·50]	9 (2·30)	(2·85)	[4·97]	24 (6·14)
Dextrality	69 (17·64)	(18·31)	[76·66]	95 (24·30)	(24·41)	[79·17]	147 (37·60)	(36·82)	[81·22]	311 (79·54)
Totals	90 (23·02)			120 (30·69)			181 (46·29)			391 (100)

(4) Grip and Refraction in "Vertical" Meridian.

Manual Laterality for Grip	Ocular Laterality for Refraction in "Vertical" Meridian									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	13 (3·31)	(3·16)	[15·48]	21 (5·34)	(4·92)	[16·03]	24 (6·11)	(6·68)	[13·48]	58 (14·76)
Ambilaterality	5 (1·28)	(1·20)	[5·95]	9 (2·29)	(1·86)	[6·87]	8 (2·03)	(2·54)	[4·49]	22 (5·60)
Dextrality	66 (16·79)	(17·02)	[78·57]	101 (25·70)	(26·55)	[77·10]	146 (37·15)	(36·07)	[82·03]	313 (79·64)
Totals	84 (21·38)			131 (33·33)			178 (45·29)			393 (100)

(5) Grip and General Astigmatism.

Manual Laterality for Grip	Ocular Laterality for General Astigmatism									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	17 (4·09)	(4·25)	[14·17]	26 (6·27)	(6·87)	[13·40]	18 (4·34)	(3·58)	[17·82]	61 (14·70)
Ambilaterality	4 (0·97)	(1·60)	[3·33]	15 (3·61)	(2·59)	[7·73]	4 (0·96)	(1·35)	[3·96]	23 (5·54)
Dextrality	99 (23·85)	(23·06)	[82·50]	153 (36·87)	(37·29)	[78·87]	79 (19·04)	(19·41)	[78·22]	331 (79·76)
Totals	120 (28·91)			194 (46·75)			101 (24·34)			415 (100)

(6) First Pull and Near Point.

Manual Laterality for First Pull	Ocular Laterality for Near Point									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	18 (4·44)	(6·14)	[27·69]*	97 (23·95)	(22·49)	[40·76]	40 (9·88)	(9·64)	[39·22]	155 (38·27)
Ambilaterality	9 (2·22)	(1·39)	[13·85]	11 (2·72)	(5·08)	[4·62]	15 (3·70)	(2·17)	[14·70]	35 (8·64)
Dextrality	38 (9·39)	(8·52)	[58·46]	130 (32·10)	(31·20)	[54·62]	47 (11·60)	(13·37)	[46·08]	215 (53·09)
Totals	65 (16·05)			238 (58·77)			102 (25·18)			405 (100)

(7) First Pull and Visual Acuity.

Manual Laterality for First Pull	Ocular Laterality for Visual Acuity									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	68 (16·46)	(16·56)	[37·57]	27 (6·54)	(6·40)	[38·58]	61 (14·77)	(14·81)	[37·65]	156 (37·77)
Ambilaterality	16 (3·87)	(4·03)	[8·84]	4 (0·97)	(1·56)	[5·71]	18 (4·36)	(3·61)	[11·11]	38 (9·20)
Dextrality	97 (23·50)	(23·24)	[53·59]	39 (9·44)	(8·99)	[55·71]	83 (20·09)	(20·80)	[51·24]	219 (53·03)
Totals	181 (43·83)			70 (16·95)			162 (39·22)			413 (100)

(8) First Pull and Refraction in "Horizontal" Meridian.

Manual Laterality for First Pull	Ocular Laterality for Refraction in "Horizontal" Meridian									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	32 (8·36)	(8·80)	[35·96]	49 (12·79)	(11·47)	[42·24]	64 (16·71)	(17·59)	[35·96]	145 (37·86)
Ambilaterality	6 (1·57)	(2·00)	[6·74]	8 (2·09)	(2·61)	[6·90]	19 (4·96)	(4·01)	[10·67]	33 (8·62)
Dextrality	51 (13·30)	(12·43)	[57·30]	59 (15·41)	(16·21)	[50·86]	95 (24·81)	(24·88)	[53·37]	205 (53·52)
Totals	89 (23·23)			116 (30·29)			178 (46·48)			383 (100)

(9) First Pull and Refraction in "Vertical" Meridian.

Manual Laterality for First Pull	Ocular Laterality for Refraction in "Vertical" Meridian									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	27 (6·99)	(8·40)	[32·14]	54 (13·99)	(12·50)	[43·20]	68 (17·62)	(17·70)	[38·42]	149 (38·60)
Ambilaterality	4 (1·04)	(1·97)	[4·76]	9 (2·33)	(2·94)	[7·20]	22 (5·70)	(4·16)	[12·43]	35 (9·07)
Dextrality	53 (13·73)	(11·39)	[63·10]	62 (16·06)	(16·94)	[49·60]	87 (22·54)	(24·00)	[49·15]	202 (52·33)
Totals	84 (21·76)			125 (32·38)			177 (45·86)			386 (100)

(10) First Pull and General Astigmatism.

Manual Laterality for First Pull	Ocular Laterality for General Astigmatism									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	44 (10·79)	(11·06)	[37·29]	76 (18·63)	(17·62)	[40·43]	36 (8·82)	(9·56)	[35·29]	156 (38·24)
Ambilaterality	10 (2·44)	(2·55)	[8·47]	19 (4·66)	(4·07)	[10·11]	7 (1·72)	(2·20)	[6·86]	36 (8·82)
Dextrality	64 (15·69)	(15·31)	[54·24]	93 (22·79)	(24·39)	[49·46]	59 (14·46)	(13·24)	[57·85]	216 (52·94)
Totals	118 (28·92)			188 (46·08)			102 (25·00)			408 (100)

(11) Last Pull and Near Point.

Manual Laterality for Last Pull	Ocular Laterality for Near Point									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	26 (6·47)	(6·89)	[40·63]	104 (25·87)	(25·52)	[43·88]	44 (10·94)	(10·87)	[43·56]	174 (43·28)
Ambilaterality	6 (1·49)	(1·78)	[9·37]	30 (7·46)	(6·60)	[12·66]	9 (2·24)	(2·81)	[8·91]	45 (11·19)
Dextrality	32 (7·96)	(7·25)	[50·00]	103 (25·63)	(26·84)	[43·46]	48 (11·94)	(11·44)	[47·53]	183 (45·53)
Totals	64 (15·92)			237 (58·96)			101 (25·12)			402 (100)

(12) Last Pull and Visual Acuity.

Manual Laterality for Last Pull	Ocular Laterality for Visual Acuity									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	83 (20·25)	(18·74)	[45·86]	24 (5·85)	(7·04)	[35·29]	67 (16·34)	(16·66)	[41·62]	174 (42·44)
Ambilaterality	22 (5·37)	(5·17)	[12·15]	6 (1·46)	(1·94)	[8·83]	20 (4·88)	(4·60)	[12·42]	48 (11·71)
Dextrality	76 (18·53)	(20·24)	[41·99]	38 (9·27)	(7·60)	[55·88]*	74 (18·05)	(18·01)	[45·96]	188 (45·85)
Totals	181 (44·15)			68 (16·58)			161 (39·27)			410 (100)

* An asterisk marks a probably significant difference in percentage value.

(13) Last Pull and Refraction in "Horizontal" Meridian.

Manual Laterality for Last Pull	Ocular Laterality for Refraction in "Horizontal" Meridian									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	39 (10·34)	(10·03)	[44·32]	38 (10·08)	(12·99)	[33·33]	85 (22·55)	(19·95)	[48·57]	162 (42·97)
Ambilaterality	11 (2·92)	(2·78)	[12·50]	18 (4·78)	(3·61)	[15·79]	16 (4·24)	(5·55)	[9·14]	45 (11·94)
Dextrality	38 (10·08)	(10·53)	[43·18]	58 (15·38)	(13·64)	[50·88]	74 (19·63)	(20·92)	[42·29]	170 (45·09)
Totals	88 (23·34)			114 (30·24)			175 (46·42)			377 (100)

(14) Last Pull and Refraction in "Vertical" Meridian.

Manual Laterality for Last Pull	Ocular Laterality for Refraction in "Vertical" Meridian									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	38 (9·90)	(9·28)	[44·71]	47 (12·24)	(13·54)	[37·90]	76 (19·79)	(19·11)	[43·43]	161 (41·93)
Ambilaterality	14 (3·65)	(2·65)	[16·47]	14 (3·64)	(3·87)	[11·29]	18 (4·69)	(5·46)	[10·28]	46 (11·98)
Dextrality	33 (8·59)	(10·21)	[38·82]	63 (16·41)	(14·88)	[50·81]	81 (21·09)	(21·00)	[46·29]	177 (46·09)
Totals	85 (22·14)			124 (32·29)			175 (45·57)			384 (100)

(15) Last Pull and General Astigmatism.

Manual Laterality for Last Pull	Ocular Laterality for General Astigmatism									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	51 (12·62)	(12·54)	[42·86]	70 (17·33)	(19·49)	[37·84]	51 (12·62)	(10·54)	[51·00]	172 (42·57)
Ambilaterality	18 (4·46)	(3·36)	[15·13]	20 (4·95)	(5·21)	[10·81]	8 (1·98)	(2·82)	[8·00]	46 (11·39)
Dextrality	50 (12·38)	(13·56)	[42·01]	95 (23·51)	(21·09)	[51·35]	41 (10·15)	(11·39)	[41·00]	186 (46·04)
Totals	119 (29·46)			185 (45·79)			100 (24·75)			404 (100)

(16) Lack of Endurance of Pull and Near Point.

Manual Laterality for Lack of Endurance of Pull	Ocular Laterality for Near Point									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	26 (6·48)	(6·73)	[40·63]	100 (24·93)	(24·69)	[42·55]	43 (10·73)	(10·72)	[42·16]	169 (42·14)
Ambilaterality	7 (1·75)	(1·35)	[10·94]	17 (4·24)	(4·97)	[7·23]	10 (2·49)	(2·16)	[9·80]	34 (8·48)
Dextrality	31 (7·73)	(7·88)	[48·43]	118 (29·43)	(28·94)	[50·22]	49 (12·22)	(12·56)	[48·04]	198 (49·38)
Totals	64 (15·96)			235 (58·60)			102 (25·44)			401 (100)

(17) Lack of Endurance of Pull and Visual Acuity.

Manual Laterality for Lack of Endurance of Pull	Ocular Laterality for Visual Acuity									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	74 (18·05)	(18·11)	[41·57]	30 (7·32)	(6·71)	[45·45]	67 (16·34)	(16·89)	[40·36]	171 (41·71)
Ambilaterality	8 (1·95)	(3·28)	[4·49]	6 (1·46)	(1·22)	[9·10]	17 (4·15)	(3·06)	[10·24]	31 (7·56)
Dextrality	96 (23·41)	(22·02)	[53·94]	30 (7·32)	(8·17)	[45·45]	82 (20·00)	(20·54)	[49·40]	208 (50·73)
Totals	178 (43·41)			66 (16·10)			166 (40·49)			410 (100)

(18) Lack of Endurance of Pull and Refraction in "Horizontal" Meridian.

Manual Laterality for Lack of Endurance of Pull	Ocular Laterality for Refraction in "Horizontal" Meridian									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	32 (8·44)	(9·19)	[36·36]	54 (14·25)	(11·91)	[47·37]	64 (16·89)	(18·48)	[36·16]	150 (39·58)
Ambilaterality	7 (1·85)	(2·14)	[7·95]	9 (2·37)	(2·78)	[7·89]	19 (5·01)	(4·31)	[10·73]	35 (9·23)
Dextrality	49 (12·93)	(11·89)	[55·69]	51 (13·46)	(15·39)	[44·74]	94 (24·80)	(23·91)	[53·11]	194 (51·19)
Totals	88 (23·22)			114 (30·08)			177 (46·70)			379 (100)

(19) Lack of Endurance of Pull and Refraction in "Vertical" Meridian.

Manual Laterality for Lack of Endurance of Pull	Ocular Laterality for Refraction in "Vertical" Meridian									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	29 (7·61)	(9·18)	[33·72]	57 (14·96)	(13·03)	[46·72]	69 (18·11)	(18·47)	[39·88]	155 (40·68)
Ambilaterality	5 (1·31)	(2·07)	[5·81]	8 (2·10)	(2·94)	[6·56]	22 (5·78)	(4·18)	[12·72]	35 (9·19)
Dextrality	52 (13·65)	(11·32)	[60·47]*	57 (14·96)	(16·05)	[46·72]	82 (21·52)	(22·76)	[47·40]	191 (50·13)
Totals	86 (22·57)			122 (32·02)			173 (45·41)			381 (100)

(20) Lack of Endurance of Pull and General Astigmatism.

Manual Laterality for Lack of Endurance of Pull	Ocular Laterality for General Astigmatism									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	55 (13·72)	(12·31)	[45·83]	79 (19·70)	(18·78)	[43·17]	31 (7·73)	(10·06)	[31·63]	165 (41·45)
Ambilaterality	12 (2·99)	(2·62)	[10·00]	18 (4·49)	(3·98)	[9·84]	5 (1·25)	(2·13)	[5·10]	35 (8·73)
Dextrality	53 (13·22)	(15·00)	[44·17]	86 (21·44)	(22·87)	[46·99]	62 (15·46)	(12·25)	[63·27]*	201 (50·12)
Totals	120 (29·93)			183 (45·63)			98 (24·44)			401 (100)

* An asterisk marks a probably significant difference in percentage value.

(21) Steadiness of Hand and Near Point.

Manual Laterality for Steadiness of Hand	Ocular Laterality for Near Point									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	25 (6·11)	(3·38)	[38·46]*	42 (10·27)	(12·59)	[17·35]	20 (4·89)	(5·30)	[19·61]	87 (21·27)
Ambilaterality	8 (1·95)	(2·10)	[12·31]	27 (6·60)	(7·81)	[11·16]	19 (4·65)	(3·29)	[18·63]	54 (13·20)
Dextrality	32 (7·83)	(10·41)	[49·23]*	173 (42·30)	(38·77)	[71·49]	63 (15·40)	(16·35)	[61·76]	268 (65·53)
Totals	65 (15·89)			242 (59·17)			102 (24·94)			409 (100)

(22) Steadiness of Hand and Visual Acuity.

Manual Laterality for Steadiness of Hand	Ocular Laterality for Visual Acuity									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	41 (9·93)	(9·13)	[22·91]	16 (3·87)	(3·52)	[23·19]	30 (7·26)	(8·41)	[18·18]	87 (21·06)
Ambilaterality	27 (6·54)	(5·87)	[15·08]	9 (2·18)	(2·27)	[13·04]	20 (4·84)	(5·42)	[12·12]	56 (13·56)
Dextrality	111 (26·87)	(28·34)	[62·01]	44 (10·66)	(10·92)	[63·77]	115 (27·85)	(26·12)	[69·70]	270 (65·38)
Totals	179 (43·34)			69 (16·71)			165 (39·95)			413 (100)

(23) Steadiness of Hand and Refraction in "Horizontal" Meridian.

Manual Laterality for Steadiness of Hand	Ocular Laterality for Refraction in "Horizontal" Meridian									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	16 (4·14)	(4·86)	[17·58]	25 (6·46)	(6·25)	[21·37]	39 (10·07)	(9·56)	[21·79]	80 (20·67)
Ambilaterality	12 (3·10)	(2·92)	[13·19]	11 (2·84)	(3·75)	[9·40]	25 (6·46)	(5·73)	[13·97]	48 (12·40)
Dextrality	63 (16·28)	(15·74)	[69·23]	81 (20·93)	(20·23)	[69·23]	115 (29·72)	(30·96)	[64·24]	259 (66·93)
Totals	91 (23·52)			117 (30·23)			179 (46·25)			387 (100)

(24) Steadiness of Hand and Refraction in "Vertical" Meridian.

Manual Laterality for Steadiness of Hand	Ocular Laterality for Refraction in "Vertical" Meridian									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	12 (3·08)	(4·55)	[14·12]	25 (6·43)	(6·74)	[19·84]	44 (11·31)	(9·53)	[24·72]	81 (20·82)
Ambilaterality	12 (3·08)	(2·81)	[14·12]	14 (3·60)	(4·16)	[11·11]	24 (6·17)	(5·88)	[13·48]	50 (12·85)
Dextrality	61 (15·69)	(14·49)	[71·76]	87 (22·36)	(21·49)	[69·05]	110 (28·28)	(30·35)	[61·80]	258 (66·33)
Totals	85 (21·85)			126 (32·39)			178 (45·76)			389 (100)

* An asterisk marks a probably significant difference in percentage value.

(25) Steadiness of Hand and General Astigmatism.

Manual Laterality for Steadiness of Hand	Ocular Laterality for General Astigmatism									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	31 (7·56)	(6·18)	[26·27]	37 (9·03)	(10·00)	[19·37]	20 (4·87)	(5·28)	[19·80]	88 (21·46)
Ambilaterality	11 (2·68)	(3·65)	[9·32]	27 (6·58)	(5·91)	[14·14]	14 (3·42)	(3·12)	[13·86]	52 (12·68)
Dextrality	76 (18·54)	(18·95)	[64·41]	127 (30·98)	(30·68)	[66·49]	67 (16·34)	(16·23)	[66·34]	270 (65·86)
Totals	118 (28·78)			191 (46·59)			101 (24·63)			410 (100)

(26) Balancing Test and Near Point.

Manual Laterality for Balancing Test	Ocular Laterality for Near Point									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	23 (5·63)	(4·84)	[34·85]	72 (17·65)	(17·88)	[29·51]	27 (6·62)	(7·18)	[27·55]	122 (29·90)
Ambilaterality	11 (2·70)	(2·89)	[16·67]	42 (10·29)	(10·70)	[17·21]	20 (4·90)	(4·30)	[20·41]	73 (17·89)
Dextrality	32 (7·85)	(8·45)	[48·48]	130 (31·86)	(31·22)	[53·28]	51 (12·50)	(12·54)	[52·04]	213 (52·21)
Totals	66 (16·18)			244 (59·80)			98 (24·02)			408 (100)

(27) Balancing Test and Visual Acuity.

Manual Laterality for Balancing Test	Ocular Laterality for Visual Acuity									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	48 (11·63)	(12·75)	[27·59]	25 (6·05)	(5·21)	[35·21]	52 (12·59)	(12·31)	[30·95]	125 (30·27)
Ambilaterality	32 (7·75)	(7·14)	[18·39]	13 (3·15)	(2·91)	[18·31]	25 (6·05)	(6·90)	[14·88]	70 (16·95)
Dextrality	94 (22·75)	(22·24)	[54·02]	33 (7·99)	(9·07)	[46·48]	91 (22·04)	(21·47)	[54·17]	218 (52·78)
Totals	174 (42·13)			71 (17·19)			168 (40·68)			413 (100)

(28) Balancing Test and Refraction in "Horizontal" Meridian.

Manual Laterality for Balancing Test	Ocular Laterality for Refraction in "Horizontal" Meridian									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	21 (5·36)	(6·90)	[22·83]	41 (10·46)	(9·05)	[33·88]	53 (13·52)	(13·39)	[29·61]	115 (29·34)
Ambilaterality	26 (6·63)	(4·42)	[28·26]	20 (5·10)	(5·83)	[16·53]	28 (7·14)	(8·62)	[15·64]	74 (18·87)
Dextrality	45 (11·48)	(12·15)	[48·91]	60 (15·31)	(15·99)	[49·59]	98 (25·00)	(23·65)	[54·75]	203 (51·79)
Totals	92 (23·47)			121 (30·87)			179 (45·66)			392 (100)

(29) Balancing Test and Refraction in "Vertical" Meridian.

Manual Laterality for Balancing Test	Ocular Laterality for Refraction in "Vertical" Meridian									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	22 (5·77)	(6·55)	[26·83]	39 (10·24)	(9·99)	[31·20]	55 (14·44)	(13·91)	[31·61]	116 (30·45)
Ambilaterality	26 (6·83)	(3·90)	[31·71]*	16 (4·20)	(5·94)	[12·80]	27 (7·08)	(8·27)	[15·52]	69 (18·11)
Dextrality	34 (8·92)	(11·07)	[41·46]	70 (18·37)	(16·88)	[56·00]	92 (24·15)	(23·49)	[52·87]	196 (51·44)
Totals	82 (21·52)			125 (32·81)			174 (45·67)			381 (100)

(30) Balancing Test and General Astigmatism.

Manual Laterality for Balancing Test	Ocular Laterality for General Astigmatism									Totals %
	Sinistrality			Ambilaterality			Dextrality			
	% of Totals		% of Column	% of Totals		% of Column	% of Totals		% of Column	
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.		
Sinistrality	34 (8·23)	(8·58)	[28·81]	62 (15·01)	(14·18)	[31·79]	28 (6·78)	(7·26)	[28·00]	124 (30·02)
Ambilaterality	27 (6·54)	(5·19)	[22·88]	28 (6·78)	(8·57)	[14·36]	20 (4·84)	(4·40)	[20·00]	75 (18·16)
Dextrality	57 (13·80)	(14·80)	[48·31]	105 (25·43)	(24·47)	[53·85]	52 (12·59)	(12·55)	[52·00]	214 (51·82)
Totals	118 (28·57)			195 (47·22)			100 (24·21)			413 (100)

* An asterisk marks a probably significant difference in percentage value.

their difference and probably not significant. There are seven exceptional cases (especially marked by an asterisk alongside such values) in which the differences are over 10·0 per cent. The largest one is 17·19 per cent. (38·46—21·27 per cent.) in the case of Steadiness test and Near Point, in the category of both manual and visual sinistrality, and this value being 3·97 times that of the probable error should be treated as significant. However, these special cases being very few in number are possibly due to the effect of random sampling or to the mixed character of certain of the tests, and they do not represent any general relation between characters in the two groups.

From the results of the comparison by the two methods just stated, we may safely assert that little relation of laterality exists between any pair of ocular and manual variates as far as judgment by percentage occurrences is concerned.

Further, for ease of comparison the actual percentages of each type of laterality given in the above table are divided in a manner similar to that used in the previous cases, namely, into three main classes, viz. same, opposite and mixed laterality. Since the two variates treated are of different groups, it will be more instructive to divide the percentages of opposite laterality into two sub-classes, viz. (1) those which are dextral in manual but sinistral in ocular variate, (2) those where the lateralities of these two variates are reversed. The category of mixed laterality may also be grouped according to the nature of the laterality in four sub-classes: (1) ocular sinistral and manual ambilateral, (2) manual sinistral and ocular ambilateral, (3) ocular dextral and manual ambilateral,

TABLE XVI.

Percentages of types of Laterality for Ocular and Manual Characters.

Pairs of Characters (Observation)	(I) Same Laterality				(II) Opposite Laterality			(III) Mixed Laterality				
								Sinistrality and Ambilaterality		Dextrality and Ambilaterality		
	Sinistrality	Ambi-laterality	Dextrality	Totals	Ocular Dextrality and Manual Sinistrality	Manual Dextrality and Ocular Sinistrality	Totals	Ocular Sinistrality and Manual Ambi-laterality	Manual Sinistrality and Ocular Ambi-laterality	Ocular Dextrality and Manual Ambi-laterality	Manual Dextrality and Ocular Ambi-laterality	Totals
Grip and Near Point (411)	3·16	3·89	20·19	27·24	2·68	12·17	14·85	0·73	8·52	1·46	47·20	57·91
„ and Visual Acuity (419)	6·21	0·71	31·74	38·66	5·72	35·08	40·80	3·10	2·63	1·68	13·13	20·54
„ and Refraction "Horizontal" (391)	3·84	2·30	37·60	43·74	6·39	17·64	24·03	1·54	4·09	2·30	24·30	32·23
„ and Refraction "Vertical" (393)	3·31	2·29	37·15	42·75	6·11	16·79	22·90	1·28	5·34	2·03	25·70	34·35
„ and General Astigmatism (415)	4·09	3·61	19·04	26·74	4·34	23·85	28·19	0·97	6·27	0·96	36·87	45·07
First Pull and Near Point (405)	4·44	2·72	11·60	18·76	9·88	9·39	19·27	2·22	23·95	3·70	32·10	61·97
„ and Visual Acuity (413)	16·46	0·97	20·09	37·52	14·77	23·50	38·27	3·87	6·54	4·36	9·44	24·21
„ and Refraction "Horizontal" (383)	8·36	2·09	24·81	35·26	16·71	13·30	30·01	1·57	12·79	4·96	15·41	34·73
„ and Refraction "Vertical" (386)	6·99	2·33	22·54	31·86	17·62	13·73	31·35	1·04	13·99	5·70	16·06	36·79
„ and General Astigmatism (408)	10·79	4·66	14·46	29·91	8·82	15·69	24·51	2·44	18·63	1·72	22·79	45·58
Last Pull and Near Point (402)	6·47	7·46	11·94	25·87	10·94	7·96	18·90	1·49	25·87	2·24	25·63	55·23
„ and Visual Acuity (410)	20·25	1·46	18·05	39·76	16·34	18·53	34·87	5·37	5·85	4·88	9·27	25·37
„ and Refraction "Horizontal" (377)	10·34	4·78	19·63	34·75	22·55	10·08	32·63	2·92	10·08	4·24	15·38	32·62
„ and Refraction "Vertical" (384)	9·90	3·64	21·09	34·63	19·79	8·59	28·38	3·65	12·24	4·69	16·41	36·99
„ and General Astigmatism (404)	12·62	4·95	10·15	27·72	12·62	12·38	25·00	4·46	17·33	1·98	23·51	47·28
Lack of Endurance of Pull (First Pull – Last Pull) and Near Point (401)	6·48	4·24	12·22	22·94	10·73	7·73	18·46	1·75	24·93	2·49	29·43	58·60
„ and Visual Acuity (410)	18·05	1·46	20·00	39·51	16·34	23·41	39·75	1·95	7·32	4·15	7·32	20·74
„ and Refraction "Horizontal" (379)	8·44	2·37	24·80	35·61	16·89	12·93	29·82	1·85	14·25	5·01	13·46	34·57
„ and Refraction "Vertical" (381)	7·61	2·10	21·52	31·23	18·11	13·65	31·76	1·31	14·96	5·78	14·96	37·01
„ and General Astigmatism (401)	13·72	4·49	15·46	33·67	7·73	13·22	20·95	2·99	19·70	1·25	21·44	45·38
Steadiness Test and Near Point (409)	6·11	6·60	15·40	28·11	4·89	7·83	12·72	1·95	10·27	4·65	42·30	59·17
„ and Visual Acuity (413)	9·93	2·18	27·85	39·96	7·26	26·87	34·13	6·54	3·87	4·84	10·66	25·91
„ and Refraction "Horizontal" (387)	4·14	2·84	29·72	36·70	10·07	16·28	26·35	3·10	6·46	6·46	20·93	36·95
„ and Refraction "Vertical" (389)	3·08	3·60	28·28	34·96	11·31	15·69	27·00	3·08	6·43	6·17	22·36	38·04
„ and General Astigmatism (410)	7·56	6·58	16·34	30·48	4·87	18·54	23·41	2·68	9·03	3·42	30·98	46·11
Balancing Test and Near Point (408)	5·63	10·29	12·50	28·42	6·62	7·85	14·47	2·70	17·65	4·90	31·86	57·11
„ and Visual Acuity (413)	11·63	3·15	22·04	36·82	12·59	22·75	35·34	7·75	6·05	6·05	7·99	27·84
„ and Refraction "Horizontal" (392)	5·36	5·10	25·00	35·46	13·52	11·48	25·00	6·63	10·46	7·14	15·31	39·54
„ and Refraction "Vertical" (381)	5·77	4·20	24·15	34·12	14·44	8·92	23·36	6·83	10·24	7·08	18·37	42·52
„ and General Astigmatism (413)	8·23	6·78	12·59	27·60	6·78	13·80	20·58	6·54	15·01	4·84	25·43	51·82
Mean	8·30 ± ·54	3·79 ± ·26	20·93 ± ·88	33·02 ± ·72	11·25 ± ·63	15·32 ± ·79	26·57 ± ·90	3·14 ± ·24	11·69 ± ·78	4·04 ± ·22	21·54 ± 1·21	40·41 ± 1·45
Standard Deviation	4·35 ± ·38	2·12 ± ·18	7·16 ± ·62	5·82 ± ·51	5·14 ± ·45	6·43 ± ·56	7·30 ± ·64	1·97 ± ·17	6·34 ± ·55	1·79 ± ·16	9·85 ± ·86	11·76 ± 1·02

(4) manual dextral and ocular ambilateral. Table XVI provides the classified percentages of each type of laterality and their mean percentages and standard deviations are also given. On an examination of this table it will be seen that the mean percentage of like laterality is 33·02 ± ·72 per cent., or only one-third of the whole population. If the mean percentage of same ambilaterality be excluded, like laterality is slightly more than the mean value of opposite laterality (2·66 per cent. more) and much less than that of mixed laterality (11·18 per cent. less). This indicates that about one person out of three will probably follow the rule of absolute laterality, right-eyedness accompanied by right-handedness and left-eyedness accompanied by left-handedness, and that the other two either may get the opposite lateralities or be ambilateral for at least one of the pair of characters. If the theory of universal laterality were accepted we should expect the highest percentages in the categories of the same dextrality and sinistrality and the lowest ones or, according to the strict theory, nothing at all in the remaining categories. Such a theory is completely out of harmony with the results of our present investigation.

Even if we could disregard two-thirds of the population and consider only the group where the lateralities of the two characters related are alike, the mean percentage of ambilaterality in this group is still significantly large, being 11·5 per cent.

In the case of opposite laterality, the individuals on the average who are dextral in hand and sinistral in vision are approximately 4·1 per cent. more numerous than those in whom these lateralities are reversed. In the case of mixed laterality the mean value of Class (4) is nearly 10·0 per cent. more than that of Class (2), while the mean value of Class (3) is only about 1·0 per cent. more than that of Class (1). In other words, when we take ocular ambilaterality there is a large variation between the percentages associated with manual dextrality and manual sinistrality, but when we take manual ambilaterality there is only a slight difference between the percentages of ocular dextrality and ocular sinistrality.

The value of the standard deviations for thirty samples of any type laterality is large and significant as in the previous cases. In other words, there is no even approximate constancy in the percentage values for the different pairs of characters.

(*b*) *On the quantitative Relationships between the Dominance of Ocular and Manual Characters.*

We have shown that there is no relationship at all in the percentage occurrences of the various classes of laterality for various pairs of ocular and manual characters. Let us consider, from the same data, if there be any quantitative relationship between them.

(1) Difference of Grip and Near Point (A_1 and B_1).

In a previous paper we have shown, on the basis of Galton's data, that there is little relation between the laterality of one manual character, Grip, and one

Standard Deviation	4·35 ± ·38	2·12 ± ·18	7·16 ± ·62	5·82 ± ·51	5·14 ± ·45	6·43 ± ·56	7·30 ± ·64	1·97 ± ·17	6·34 ± ·55	1·79 ± ·16	9·85 ± ·86	11·76 ± 1·02

ocular character, Visual Acuity, measured by a test of reading diamond type. We have now to investigate the quantitative relation, if any, between dominance of Grip and dominance of five various visual characters. We shall first deal with the difference between the Grip of the two hands in pounds and between the

TABLE XVII.

Laterality of Strength of Grip and Distance of Near Point.

Difference of Distance of Near Point in Inches. (Central Values.) Right minus Left.

Difference of Strength of Grip in Pounds. (Central Values.) Right minus Left.	−6·50	−6·00	−5·50	..	−4·00	..	−2·00	−1·50	−1·00	−0·50	0	0·50	1·00	1·50	2·00	2·50	3·00	..	5·50	
−15	—	—	—	..	—	..	—	—	—	—	1	—	—	—	—	—	—	..	—	1
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
−12	—	—	—	..	—	..	—	—	—	—	1	—	—	—	—	—	—	..	—	1
−11	—	—	—	..	—	..	—	—	—	—	—	—	—	—	—	—	—	..	—	0
−10	—	—	—	..	—	..	—	—	—	—	2	—	—	—	—	—	—	..	—	2
− 9	—	—	—	..	—	..	—	—	—	—	2	—	—	—	—	—	—	..	—	2
− 8	—	—	—	..	—	..	—	—	—	—	4	1	—	—	—	—	—	..	—	5
− 7	—	—	—	..	—	..	—	—	1	—	—	1	—	—	—	—	—	..	—	2
− 6	—	—	—	..	—	..	—	—	—	—	—	—	1	—	—	—	—	..	—	1
− 5	—	—	—	..	—	..	—	—	—	—	3	—	—	—	—	—	—	..	—	3
− 4	—	—	—	..	—	..	—	—	—	—	7	—	—	—	—	—	—	..	—	7
− 3	—	—	—	..	1	..	—	—	1	4	2	2	—	—	1	—	—	..	—	11
− 2	—	—	—	..	—	..	—	—	1	1	6	2	2	—	—	—	—	..	—	12
− 1	—	—	—	..	—	..	—	—	2	2	7	—	1	—	—	—	—	..	—	12
0	—	—	—	..	—	..	—	1	1	1	16	3	1	1	—	—	—	..	1	25
1	—	—	—	..	—	..	—	—	1	5	18	6	1	—	1	—	—	..	—	32
2	—	—	—	..	1	..	—	—	3	1	17	3	2	3	—	—	—	..	—	30
3	—	—	—	..	—	..	1	—	—	4	13	5	1	—	—	1	—	..	—	25
4	1	—	—	..	—	..	—	1	1	3	19	3	2	—	—	—	—	..	—	30
5	—	—	—	..	—	..	—	—	2	4	23	3	—	1	1	—	—	..	—	34
6	—	—	—	..	—	..	—	—	—	2	20	5	3	—	—	—	—	..	—	30
7	—	—	—	..	—	..	—	—	1	2	15	4	1	1	—	—	—	..	—	24
8	—	—	1	..	—	..	—	—	—	3	25	5	2	2	—	—	1	..	—	39
9	—	—	—	..	—	..	—	—	—	1	10	—	1	—	1	—	—	..	—	13
10	—	—	—	..	—	..	—	—	1	1	9	5	1	—	—	—	1	..	—	18
11	—	—	—	..	—	..	—	—	—	4	8	1	1	—	—	—	—	..	—	14
12	—	—	—	..	—	..	—	—	2	2	4	4	—	—	—	—	1	..	—	13
13	—	—	—	..	—	..	—	—	—	—	3	3	—	1	—	—	—	..	—	7
14	—	—	—	..	—	..	1	—	—	—	3	1	1	—	—	—	—	..	—	6
15	—	—	—	..	—	..	—	—	—	1	1	1	—	—	—	—	—	..	—	3
16	—	—	—	..	—	..	—	—	—	—	2	1	—	—	—	—	—	..	—	3
17	—	—	—	..	—	..	—	—	—	—	1	1	—	—	—	—	—	..	—	2
18	—	—	—	..	—	..	—	—	—	—	—	—	—	—	—	—	—	..	—	0
19	—	—	—	..	—	..	—	—	—	—	1	—	—	—	—	—	—	..	—	1
20	—	—	—	..	—	..	—	—	—	—	—	—	—	—	—	—	—	..	—	0
21	—	—	—	..	—	..	—	—	—	—	1	1	—	—	—	—	—	..	—	2
22	—	—	—	..	—	..	—	—	—	—	1	—	—	—	—	—	—	..	—	1
	1	0	1	..	2	..	2	2	17	41	245	61	21	9	4	1	3	..	1	411

distance of the Near Point of the two eyes in inches. Table XVII gives the data on which the following numerical results are based. They are taken from 411 cases with the exception of a few incomplete results. The correlation constants are:

Product moment correlation coefficient: $r_{A_1B_1} = \cdot0575 \pm \cdot0332$.

Correlation ratios:

Laterality of Grip on Laterality of Near Point:

$$\eta^2_{A_1B_1} = \cdot 014125, \quad \bar{\eta}^2_{A_1B_1} = \cdot 012165 \pm \cdot 005152.$$

Laterality of Near Point on Laterality of Grip:

$$\eta^2_{B_1A_1} = \cdot 025049, \quad \bar{\eta}^2_{B_1A_1} = \cdot 034063 \pm \cdot 008524.$$

It is quite clear that the value of $r_{A_1B_1}$, i.e. the correlation coefficient for the given two variates, is too small to be significant. Both correlation ratios are also insignificant considering the value of $\bar{\eta}^2$. Diagram IX shows the two regression lines which, considering the indicated values of the standard deviation, do not differ significantly from the mean horizontals.

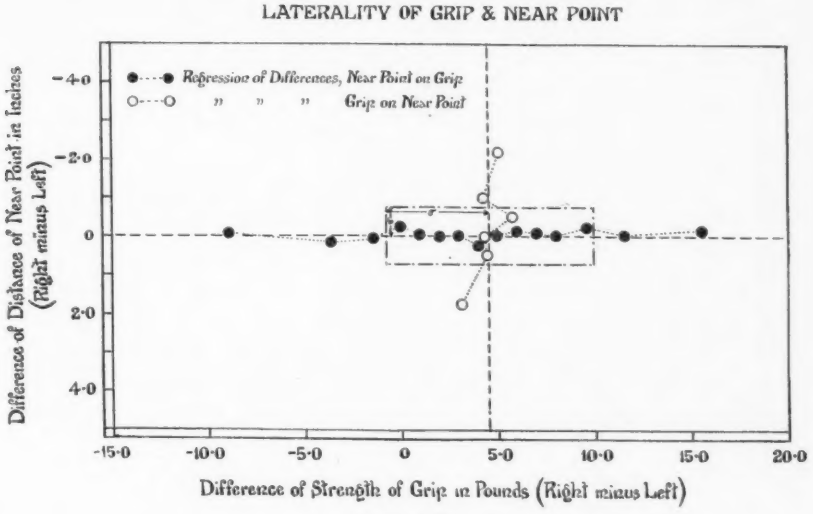


Diagram IX.

(2) Difference of Grip and Visual Acuity (A_1 and B_2).

As we have already stated the Visual Acuity was measured by the Snellen Scale and not by the diamond type used by Galton. We have $r_{A_1B_2} = \cdot 0194 \pm \cdot 0330$, calculated from Table XVIII. This value is quite insignificant with regard to its probable error, therefore dominance of grip in right hand is not associated with dominance of visual acuity in right eye. Applying the test of the correlation ratio the results are as follows:

$$\eta^2_{A_1B_2} = \cdot 025313, \quad \bar{\eta}^2_{A_1B_2} = \cdot 028640 \pm \cdot 007763,$$

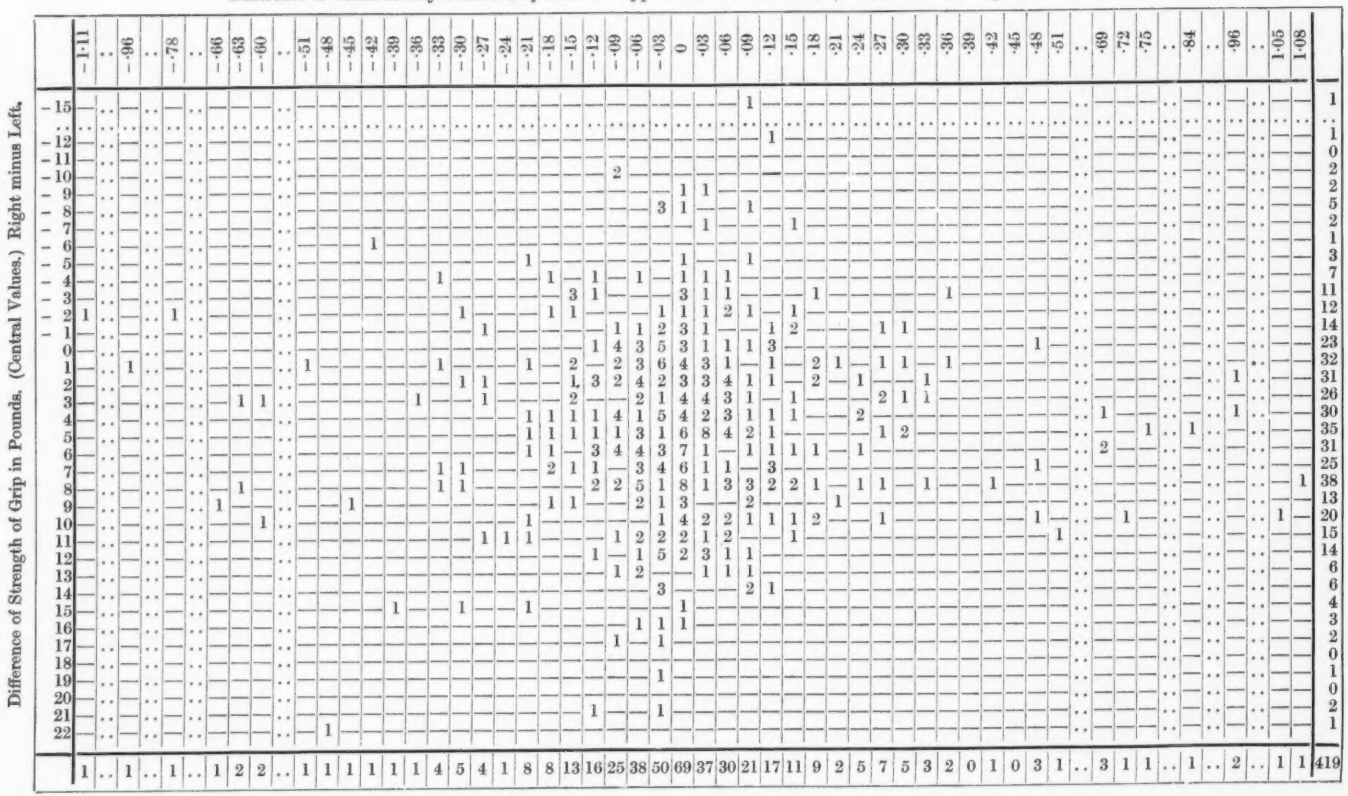
$$\eta^2_{B_2A_1} = \cdot 033547, \quad \bar{\eta}^2_{B_2A_1} = \cdot 033413 \pm \cdot 008364.$$

Both values of $\eta^2 - \bar{\eta}^2$ are very small and neither is, therefore, significant. The two regression lines in Diagram X are sensibly horizontal as in the previous case.

TABLE XVIII.

Laterality of Strength of Grip and Visual Acuity.

Difference of Visual Acuity measured by Snellen's Type reduced to decimals. (Central Values.) Right minus Left.



It may be noted that these results agree very well with those from Galton's data in spite of the fact that different subjects were tested and that the Visual Acuity was measured by another method.

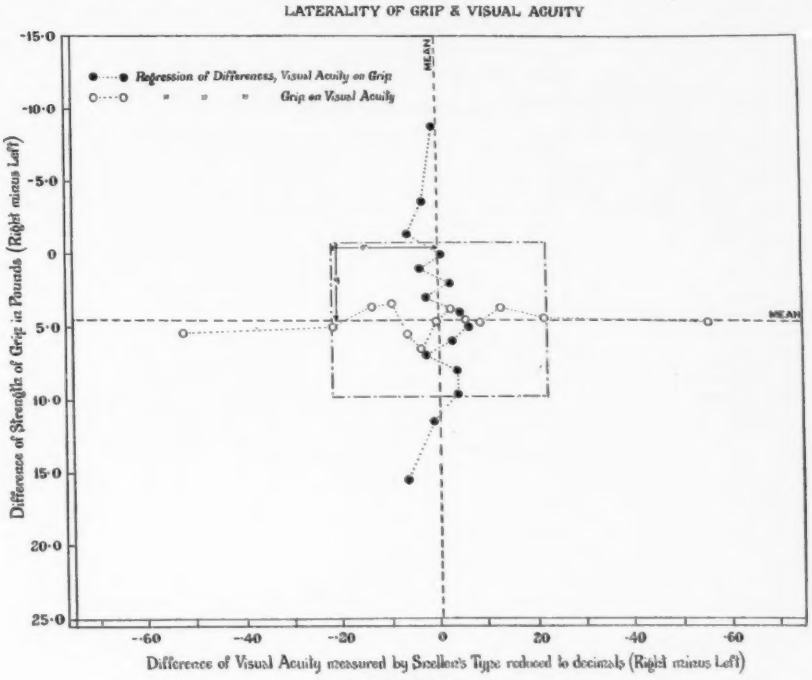


Diagram X.

(3) Difference of Grip and Refractions (A_1 and B_3, B_4).

The dominance of Grip was tested against the Refractions measured in dioptres both for the vertical and the horizontal meridians. The data have been separately tabulated in Tables XIX and XX (pp. 132—3). We have the following results:

$$r_{A_1B_3} = \cdot0114 \pm \cdot0341 \text{ (391 cases)}, \quad \begin{cases} \eta^2_{A_1B_3} = \cdot007694, & \overline{\eta}^2_{A_1B_3} = \cdot020460 \pm \cdot006821, \\ \eta^2_{B_3A_1} = \cdot022954, & \overline{\eta}^2_{B_3A_1} = \cdot033248 \pm \cdot008638, \end{cases}$$

$$r_{A_1B_4} = -\cdot0097 \pm \cdot0340 \text{ (393 cases)}, \quad \begin{cases} \eta^2_{A_1B_4} = \cdot015204, & \overline{\eta}^2_{A_1B_4} = \cdot020356 \pm \cdot006786, \\ \eta^2_{B_4A_1} = \cdot016139, & \overline{\eta}^2_{B_4A_1} = \cdot033079 \pm \cdot008594. \end{cases}$$

Obviously none of the above constants can be regarded as of significant value. The two product moment correlation coefficients have opposite signs but the values are smaller than their probable errors and accordingly insignificant. All the correlation ratios are also insignificant. From these results we may assert that the dominance of Grip of the hands has no relation at all to the dominance of the two Refractions of the eyes. Diagrams XI and XII show the regression lines for each pair of variates. It will be seen clearly that we could hardly do anything better than use the mean lines as regression lines.

TABLE XIX.

Laterality of Grip and Refraction in the Horizontal Meridian.

Difference of Refraction in the Horizontal Meridian in Dioptres. (Central Values.) Right minus Left.

Difference of Strength of Grip in Pounds. (Central Values.) Right minus Left.

	+3·50	..	+2·75	..	+2·00	+1·75	+1·50	+1·25	+1·00	+0·75	+0·50	+0·25	0	−0·25	−0·50	−0·75	−1·00	−1·25	−1·50	−1·75	−2·00	−2·25	−2·50	−2·75	−3·00	−3·25	−3·50	..	−4·25	−4·50	
−15	—	..	—	..	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	1
−12	—	..	—	..	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	..	—	—	1
−11	—	..	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	0
−10	—	..	—	..	—	—	—	—	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	2
−9	—	..	—	..	—	—	—	—	—	—	—	—	1	—	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	2
−8	—	..	—	..	—	—	—	—	1	—	—	1	1	1	1	1	—	—	—	—	—	—	—	—	—	—	—	..	—	—	6
−7	—	..	—	..	—	—	—	—	—	—	—	—	1	—	—	1	—	—	—	—	—	—	—	—	—	—	—	..	—	—	2
−6	—	..	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	0
−5	—	..	—	..	—	—	—	—	—	—	—	—	1	1	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	3
−4	—	..	—	..	—	—	—	—	—	—	2	—	1	—	1	1	—	—	—	—	—	—	—	—	—	—	—	..	—	—	5
−3	—	..	—	..	—	—	—	—	—	—	1	1	3	3	—	1	—	—	—	—	—	—	—	—	—	—	—	..	—	—	9
−2	—	..	—	..	—	—	1	1	—	—	1	2	3	1	1	1	—	—	—	—	—	—	—	—	—	—	—	..	1	—	12
−1	—	..	—	..	—	—	—	—	—	1	—	1	4	4	1	—	1	1	—	—	—	—	—	—	—	—	—	..	—	—	13
0	—	..	—	..	—	—	—	—	—	1	2	3	9	6	3	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	24
1	—	..	—	..	—	—	—	—	—	1	—	3	11	3	6	4	—	—	—	—	1	—	—	—	1	—	—	..	—	—	30
2	—	..	—	..	—	—	1	—	2	—	3	3	4	4	7	—	1	1	1	—	1	—	—	—	1	—	—	..	—	—	29
3	—	..	—	..	—	—	—	—	—	—	2	2	13	3	2	2	—	—	1	—	—	—	—	—	—	—	—	..	—	—	25
4	—	..	—	..	—	—	—	1	—	1	2	2	8	4	3	2	—	1	2	—	—	—	—	—	1	—	—	..	—	—	27
5	—	..	—	..	—	—	—	—	—	—	—	5	8	6	7	1	—	—	—	1	—	—	—	—	1	—	—	..	—	—	29
6	—	..	—	..	—	—	—	—	—	1	4	—	13	2	4	5	—	—	—	—	—	—	—	—	—	—	—	..	—	—	29
7	—	..	—	..	—	—	—	—	1	—	—	2	6	6	6	—	—	—	1	—	—	—	1	—	—	1	—	..	—	—	24
8	—	..	1	..	—	1	—	—	—	3	—	5	11	2	4	4	4	—	1	—	1	—	—	1	1	—	—	..	—	—	39
9	1	..	—	..	—	—	1	—	—	—	—	4	1	2	2	1	—	1	—	—	—	—	—	—	—	—	—	..	—	—	13
10	—	..	—	..	—	—	—	—	1	—	—	2	3	3	3	1	—	2	—	—	—	—	—	—	—	—	1	..	—	—	16
11	—	..	—	..	1	—	—	—	—	1	—	1	4	5	3	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	15
12	—	..	—	..	—	—	1	—	—	—	2	—	6	3	—	—	1	—	—	—	—	—	—	—	—	—	—	..	—	1	14
13	—	..	—	..	—	—	—	—	—	1	2	1	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	..	—	—	5
14	—	..	—	..	—	—	—	—	—	1	—	2	1	—	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	5
15	—	..	—	..	—	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	2
16	—	..	—	..	—	—	—	—	—	—	—	—	2	—	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	3
17	—	..	—	..	—	—	—	—	—	—	—	—	1	—	—	1	—	—	—	—	—	—	—	—	—	—	—	..	—	—	2
18	—	..	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	0
19	—	..	—	..	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	1
20	—	..	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	0
21	—	..	—	..	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	2
22	—	..	—	..	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	1
	1	..	1	..	1	1	4	2	7	10	22	41	120	60	59	28	7	6	6	1	3	0	1	1	5	1	1	..	1	1	391

Difference of Strength of Grip in Pounds. (Central Values.) Right minus Left.

TABLE XX.

Laterality of Grip and Refraction in the Vertical Meridian.

Difference of Refraction in the Vertical Meridian in Dioptres. (Central Values.) Right minus Left.

Difference of Strength of Grip in Pounds. (Central Values.) Right minus Left.

	+3·00	+2·75	+2·50	+2·25	+2·00	+1·75	+1·50	+1·25	+1·00	+0·75	+0·50	+0·25	0	−0·25	−0·50	−0·75	−1·00	−1·25	−1·50	−1·75	−2·00	−2·25	−2·50	−2·75	−3·00	−3·25	−3·50	..	−4·50	
−15	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	1
−12	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	..	—	1
−11	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	0
−10	—	—	—	—	—	—	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	2
−9	—	—	—	—	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	2
−8	—	—	—	—	—	—	—	—	—	—	—	—	4	1	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	6
−7	—	—	—	—	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	2
−6	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	..	—	1
−5	—	—	—	—	—	—	—	—	—	—	—	—	1	1	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	3
−4	—	—	—	—	—	—	—	—	—	—	2	—	2	2	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	6
−3	—	—	—	—	—	—	—	—	—	—	1	—	4	4	—	—	—	—	2	—	—	—	—	—	—	—	—	..	—	11
−2	—	—	—	—	—	—	—	—	—	2	—	2	3	1	—	1	—	1	—	—	—	1	—	—	—	—	—	..	—	11
−1	—	—	—	—	—	—	—	—	—	—	1	1	4	2	3	1	—	—	—	—	—	—	—	—	—	—	—	..	—	12
0	—	—	—	—	—	—	—	—	—	—	2	3	9	6	2	—	—	—	—	—	—	—	—	—	—	—	—	..	—	22
1	—	—	—	—	—	—	—	1	—	2	1	1	12	5	4	1	1	—	—	1	—	—	—	—	—	—	1	..	—	30
2	—	—	—	—	—	—	2	—	1	—	—	2	10	7	4	1	—	—	1	—	—	—	1	—	1	—	—	..	—	30
3	—	—	—	—	—	—	1	—	—	—	—	—	11	7	1	2	—	1	—	1	—	—	—	—	1	—	—	..	—	25
4	—	—	1	—	1	—	—	—	2	—	2	3	9	3	3	2	1	—	—	—	—	—	—	—	—	—	—	..	1	28
5	—	—	—	—	—	—	—	—	1	—	1	3	10	5	2	5	1	—	—	—	—	—	—	—	—	1	—	..	—	29
6	—	—	—	—	—	—	—	—	—	—	2	4	10	9	2	2	—	—	—	—	—	—	—	—	—	—	—	..	—	29
7	—	—	—	—	—	—	—	—	1	1	—	2	8	5	3	—	—	—	—	—	2	1	—	—	—	—	—	..	—	23
8	—	1	1	—	1	1	—	2	—	1	1	3	11	5	2	3	2	1	1	—	—	—	—	1	1	—	—	..	—	38
9	1	—	—	—	—	—	1	—	—	—	1	—	3	3	1	2	1	—	—	—	—	—	—	—	—	—	—	..	—	13
10	—	—	—	—	—	—	—	—	—	—	—	3	4	1	4	1	1	1	—	—	—	—	—	—	—	—	1	..	—	16
11	—	—	—	—	1	—	—	—	—	—	1	1	4	4	3	1	—	—	—	—	—	—	—	—	—	—	—	..	—	15
12	—	—	1	—	—	—	—	—	1	—	2	1	2	2	1	—	1	—	—	—	2	—	—	—	—	—	—	..	—	13
13	—	—	—	—	—	—	1	—	—	—	1	—	—	2	2	1	—	—	—	—	—	—	—	—	—	—	—	..	—	7
14	—	—	—	—	—	—	—	—	—	—	1	1	1	1	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	5
15	—	—	—	—	—	—	—	—	—	—	—	1	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	3
16	—	—	—	—	—	—	—	—	—	—	—	—	2	—	—	—	1	—	—	—	—	—	—	—	—	—	—	..	—	3
17	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	1	—	—	—	—	—	—	—	—	—	—	—	..	—	2
18	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	0
19	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	1
20	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	0
21	—	—	—	—	—	—	—	—	—	—	—	—	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	2
22	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	1
	1	1	3	0	3	1	8	3	6	6	19	33	131	77	41	25	9	5	4	2	4	2	1	1	3	1	2	..	1	393

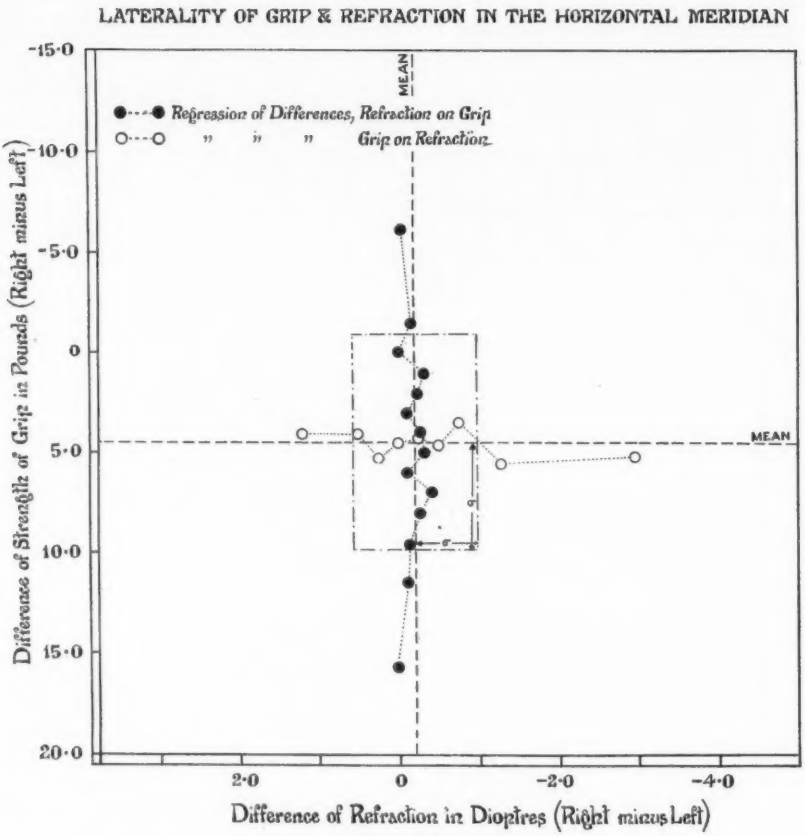

LATERALITY OF GRIP & REFRACTION IN THE HORIZONTAL MERIDIAN
MEAN
Regression of Differences, Refraction on Grip
" " " Grip on Refraction
MEAN
-15·0
-10·0
-5·0
0
5·0
10·0
15·0
20·0
2·0
0
-2·0
-4·0
Difference of Strength of Grip in Pounds (Right minus Left)
Difference of Refraction in Dioptres (Right minus Left)

Diagram XI.

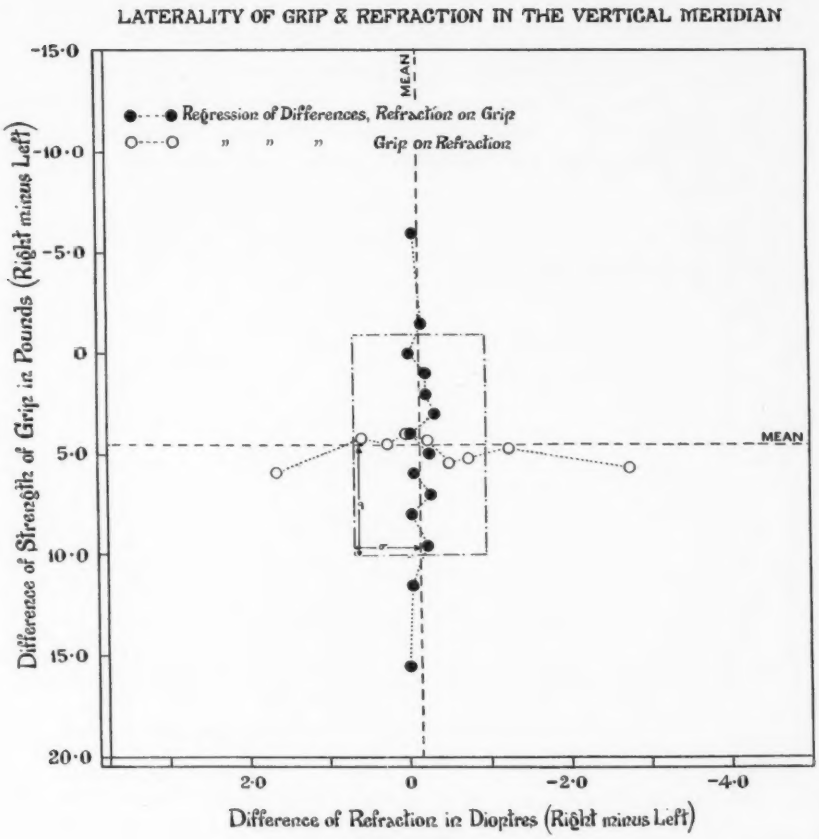

LATERALITY OF GRIP & REFRACTION IN THE VERTICAL MERIDIAN
MEAN
Regression of Differences, Refraction on Grip
" " " Grip on Refraction
MEAN
-15·0
-10·0
-5·0
0
5·0
10·0
15·0
20·0
Difference of Strength of Grip in Pounds (Right minus Left)
2·0
0
-2·0
-4·0
Difference of Refraction in Dioptres (Right minus Left)

Diagram XII.

(4) Difference of Grip and General Astigmatism (A_1 and B_5).

The last ocular character to be considered in association with Grip is General Astigmatism. The data from 415 cases are presented in Table XXI. This leads to the following value: $r_{A_1B_5} = -\cdot0544 \pm \cdot0330$. It has a *negative* sign and is a little higher than r in the case of Vertical Refraction with Grip, but as it is only 1·3 times its probable error, it is of no significance. The correlation ratios are:

$$\eta^2_{A_1B_5} = \cdot010156, \quad \bar{\eta}^2_{A_1B_5} = \cdot014458 \pm \cdot005583,$$
$$\eta^2_{B_5A_1} = \cdot035167, \quad \bar{\eta}^2_{B_5A_1} = \cdot031325 \pm \cdot008147.$$

There are no significant differences between η^2 and $\bar{\eta}^2$. The two observation regression lines of A_1 on B_5 and B_5 on A_1 are plotted graphically in Diagram XIII; they show no significant deviations from the mean horizontals.

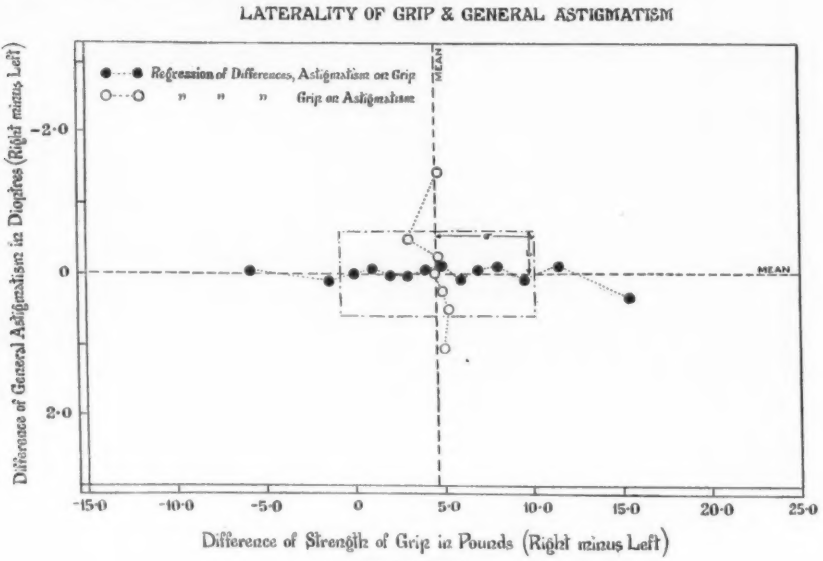


Diagram XIII.

From the results given above we may safely conclude that superiority of Grip on one side is in no way associated with superiority of any of the ocular characters tested by us.

Our next step should be to study the relations of other manual characters with each of the above ocular characters. Owing to the limited space at our disposal we have not published the other twenty-five correlation tables and the diagrams—the character of these is adequately illustrated on Grip—but the results for r and η^2 for the remaining four manual characters and each of the ocular characters are given in Table XXII (p. 138). It will be seen that all the values of r are very small and insignificant in comparison with their probable errors. For instance the

TABLE XXI.

Laterality of Grip and General Astigmatism.

Difference of General Astigmatism in Dioptres. (Central Values.) Right minus Left.

Difference of Strength of Grip in Pounds. (Central Values.) Right minus Left.

	+3·50	..	+2·00	+1·75	+1·50	+1·25	+1·00	+0·75	+0·50	+0·25	0	−0·25	−0·50	−0·75	−1·00	−1·25	−1·50	−1·75	−2·00	−2·25	−2·50	−2·75	−3·00	..	−4·00	−4·25	−4·50	
−15	—	..	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	—	1
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
−12	—	..	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	—	1
−11	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	—	0
−10	—	..	—	—	—	—	—	—	—	—	—	—	2	—	—	—	—	—	—	—	—	—	—	..	—	—	—	2
−9	—	..	—	—	—	—	—	1	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	—	2
−8	—	..	—	—	—	—	—	—	—	1	3	—	—	1	1	—	—	—	—	—	—	—	—	..	—	—	—	6
−7	—	..	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	—	2
−6	—	..	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	—	1
−5	—	..	—	—	—	—	—	—	—	—	3	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	—	3
−4	—	..	—	—	—	—	—	—	—	1	2	—	4	—	—	—	—	—	—	—	—	—	—	..	—	—	—	7
−3	—	..	—	—	—	—	—	1	1	2	6	1	—	—	—	—	—	—	—	—	—	—	—	..	—	—	—	11
−2	—	..	1	—	—	1	—	2	2	—	5	1	—	—	—	1	—	—	—	—	—	—	—	..	—	—	—	13
−1	—	..	—	—	—	—	—	—	1	2	2	2	5	—	—	—	—	—	—	—	—	—	—	..	—	—	—	12
0	—	..	—	—	—	—	—	1	2	1	15	2	1	—	1	—	—	—	—	—	—	—	—	..	—	—	—	23
1	—	..	—	—	—	—	—	1	5	1	15	3	4	1	—	—	—	—	—	—	1	—	—	..	—	—	—	31
2	—	..	—	—	—	—	—	3	2	4	14	3	4	1	—	—	—	—	—	—	—	—	—	..	—	—	—	31
3	—	..	—	—	—	—	—	2	2	6	15	—	—	—	—	—	—	—	—	—	—	—	1	..	—	—	—	26
4	—	..	—	—	—	2	—	3	2	3	13	1	1	1	1	1	—	—	—	—	—	—	—	..	—	—	1	29
5	—	..	—	—	—	—	1	—	5	4	12	4	2	1	1	1	1	—	1	—	—	—	—	..	—	—	—	33
6	—	..	—	—	—	—	1	—	4	5	14	4	1	—	1	—	—	—	—	—	—	—	—	..	—	—	—	30
7	—	..	—	—	—	—	—	1	1	2	12	5	4	—	—	—	—	—	—	—	—	—	—	..	—	—	—	25
8	—	..	—	—	—	—	—	1	6	4	16	3	3	2	2	1	—	—	1	—	—	—	—	..	—	—	—	39
9	—	..	1	—	—	—	—	—	1	1	6	1	4	—	—	—	—	—	—	—	—	—	—	..	—	—	—	14
10	—	..	—	—	—	—	—	1	4	—	11	2	—	1	—	—	—	—	—	—	—	—	—	..	—	—	—	19
11	—	..	—	—	—	—	—	2	—	2	7	1	2	1	—	—	—	—	—	—	—	—	—	..	—	—	—	15
12	—	..	—	—	—	1	—	—	1	4	5	—	—	—	1	—	1	—	—	—	—	—	—	..	1	—	—	14
13	—	..	—	—	—	—	1	—	1	1	2	1	1	—	—	—	—	—	—	—	—	—	—	..	—	—	—	7
14	1	..	—	—	—	—	—	—	1	1	3	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	—	6
15	—	..	—	—	—	—	1	—	—	—	2	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	—	3
16	—	..	—	—	—	—	—	—	1	—	2	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	—	3
17	—	..	—	—	—	—	—	—	—	—	2	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	—	2
18	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	—	0
19	—	..	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	—	1
20	—	..	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	—	0
21	—	..	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	—	2
22	—	..	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	..	—	—	—	1
	1	..	2	0	1	4	4	19	42	47	194	34	38	9	8	4	2	0	2	0	1	0	1	..	1	0	1	415

TABLE XXII.

Correlation Constants for Laterality of Ocular and Manual Characters.

Ocular and Manual Characters (Cases)	Product Moment Correlation Coefficients	Numbers of Occurrences in 30 samples at each value of observed r if expected Mean $r=0$	Correlation Ratios of Ocular Character on Manual Character		Correlation Ratios of Manual Character on Ocular Character	
			Values of $\eta^2_{B\cdot A}$	Values of $\bar{\eta}^2_{B\cdot A}$ and their Probable Errors	Values of $\eta^2_{A\cdot B}$	Values of $\bar{\eta}^2_{A\cdot B}$ and their Probable Errors
Grip and Near Point ... ... (411)	$r_{A_1B_1}=$ ·0575±·0332	7·50	·014125	·012165±·005152	·025049	·034063±·008524
Grip and Visual Acuity ... ... (419)	$r_{A_1B_2}=$ ·0194±·0330	20·88	·025313	·028640±·007763	·033547	·033413±·008364
Grip and Refraction in the "Horizontal" Meridian ... (391)	$r_{A_1B_3}=$ ·0114±·0341	24·54	·007674	·020460±·006821	·022954	·033248±·008638
Grip and Refraction in the "Vertical" Meridian ... (393)	$r_{A_1B_4}=-$·0097±·0340	25·50	·015204	·020356±·006786	·016139	·033079±·008594
Grip and General Astigmatism ... ... (415)	$r_{A_1B_5}=-$·0544±·0330	8·28	·010156	·014458±·005583	·035167	·031325±·008147
First Pull and Near Point ... ... (405)	$r_{A_2B_1}=-$·0354±·0335	14·34	·011905	·012346±·005227	·012871	·032099±·008343
First Pull and Visual Acuity ... ... (413)	$r_{A_2B_2}=$ ·0211±·0332	20·22	·015314	·026634±·007548	·035261	·031477±·008186
First Pull and Refraction in the "Horizontal" Meridian (383)	$r_{A_2B_3}=$ ·0010±·0345	29·52	·023532	·020888±·006961	·024386	·031332±·008480
First Pull and Refraction in the "Vertical" Meridian (386)	$r_{A_2B_4}=-$·0415±·0343	12·18	·018703	·020725±·006908	·020435	·031088±·008416
First Pull and General Astigmatism ... ... (408)	$r_{A_2B_5}=-$·0031±·0334	28·56	·004655	·014670±·005664	·028268	·031785±·008264
Last Pull and Near Point ... ... (402)	$r_{A_3B_1}=-$·0589±·0355	7·14	·006520	·012438±·005266	·028063	·027363±·007752
Last Pull and Visual Acuity ... ... (410)	$r_{A_3B_2}=+$·0517±·0332	9·12	·040654	·029268±·007931	·031394	·029268±·007931
Last Pull and Refraction in the "Horizontal" Meridian (377)	$r_{A_3B_3}=-$·0327±·0347	15·48	·011250	·018568±·006622	·028594	·031830±·008612
Last Pull and Refraction in the "Vertical" Meridian ... (384)	$r_{A_3B_4}=-$·0461±·0343	10·74	·014593	·018229±·006504	·030509	·028646±·008110
Last Pull and General Astigmatism ... ... (404)	$r_{A_3B_5}=$ ·0067±·0336	26·88	·006323	·014851±·005734	·038903	·029703±·008048
Lack of Endurance of Pull and Near Point ... ... (401)	$r_{A_4B_1}=$ ·0363±·0336	13·98	·003114	·012469±·005279	·045930	·034913±·008733
,, ,, ,, and Visual Acuity ... (410)	$r_{A_4B_2}=-$·0323±·0333	15·48	·022811	·029268±·007931	·047621	·031707±·008244
,, ,, ,, and Refraction "Horizontal" (379)	$r_{A_4B_3}=-$·0520±·0346	8·94	·023057	·018470±·006588	·027934	·034301±·008906
,, ,, ,, and Refraction "Vertical" ... (381)	$r_{A_4B_4}=$ ·0497±·0345	9·66	·028146	·018373±·006554	·015625	·034121±·008860
,, ,, ,, and General Astigmatism ... (401)	$r_{A_4B_5}=-$·0043±·0337	27·84	·008542	·014963±·005776	·041302	·032419±·008426
Steadiness of Hand and Near Point ... ... (409)	$r_{A_5B_1}=$ ·0250±·0333	18·54	·023751	·012225±·005177	·034195	·019560±·006523
Steadiness of Hand and Visual Acuity ... ... (413)	$r_{A_5B_2}=$ ·0477±·0331	6·84	·036207	·029056±·007874	·029256	·021792±·006845
Steadiness of Hand and Refraction "Horizontal" ... (387)	$r_{A_5B_3}=-$·0232±·0343	19·38	·015953	·018088±·006453	·018820	·023256±·007297
Steadiness of Hand and Refraction "Vertical" ... (387)	$r_{A_5B_4}=-$·0210±·0342	20·22	·018404	·020566±·006854	·016058	·023136±·007261
Steadiness of Hand and General Astigmatism ... (410)	$r_{A_5B_5}=$ ·0316±·0333	15·84	·021076	·014634±·005650	·020990	·021951±·006894
Balancing Test and Near Point ... ... (408)	$r_{A_6B_1}=$ ·0443±·0333	11·22	·007370	·012255±·005189	·023234	·022059±·006927
Balancing Test and Visual Acuity ... ... (413)	$r_{A_6B_2}=$ ·0445±·0331	11·22	·038554	·029056±·007874	·011549	·021792±·006845
Balancing Test and Refraction "Horizontal" ... (392)	$r_{A_6B_3}=-$·0190±·0341	21·12	·028361	·020408±·006803	·026817	·022959±·007206
Balancing Test and Refraction "Vertical" ... ... (381)	$r_{A_6B_4}=-$·0284±·0345	17·04	·026507	·020997±·006997	·020193	·023622±·007412
Balancing Test and General Astigmatism ... ... (413)	$r_{A_6B_5}=$ ·0498±·0331	9·54	·014636	·014528±·005610	·039567	·021792±·006845
Mean	$r_{A\cdot B}=$ ·0012±·0045					
Standard Deviation	$\sigma_r=$ ·0365±·0032					

highest value we have obtained, i.e. $r_{A_1B_1} = \cdot0575 \pm \cdot0332$, is only 1·7 times its probable error. The coefficients vary from $-\cdot0589$ to $+\cdot0575$, half with the positive sign and the other half with the negative sign. If we group these coefficients on a rough scale we obtain the following frequencies:

	$r<-\cdot050$	$-\cdot050<r<-\cdot025$	$-\cdot025<r<0$	$0<r<+\cdot025$	$+\cdot025<r<+\cdot050$	$+\cdot050<r$
Observed	3	6	6	5·5	7·5	2
Expected	4·8	4·5	5·7	5·7	4·5	4·8

It is very interesting to see that the coefficients seem to be symmetrically distributed, each lying within the limit of its probable error around the central value zero. Although we are unable to represent the distribution of correlation coefficients for all possible ocular and manual characters, we have 30 samples of such from nearly 400 individuals, and accordingly a distribution like the above hardly seems to be accidental. The observed mean $r = +\cdot0012 \pm \cdot0045$, and the observed standard deviation of the samples $= \cdot0365 \pm \cdot0032$. If we assume the expected mean $r = 0$, then in case there were no correlation the standard deviation of random samples from the 400 cases $= \dfrac{1}{\sqrt{400}} = \cdot05$, and the probable error of the expected mean r becomes ·0061. In other words:

Expected mean $r = 0 \pm \cdot0061$, Expected S.D. $= \cdot0500 \pm \cdot0044$.

Observed mean $r = \cdot0012 \pm \cdot0045$, Observed S.D. $= \cdot0365 \pm \cdot0032$.

We can see from these results that, if the probable errors are taken into account, the differences between the observed quantities and those expected on the hypothesis of independence are quite insignificant. There is little doubt that these correlation coefficients can reasonably flow from a population in which there is no association between manual and ocular laterality. We can test this matter in another way: here are 30 samples of about 400 individuals; let us ask, could the population r be zero on the basis of the observed coefficients? The values given in the third column of Table XXII are found by the area of $\frac{1}{2}(1-\alpha)$ from the Table of the Probability Integral. They indicate the probable number of occurrences in 30 random samples which would lie outside the range $\pm$ observed r. No single r has any high degree of improbability. Thus in more than half the cases an r as great or greater would occur over 15 times in 30. Further, turning back to the rough frequency grouping above, the $\chi^2 = 4\cdot84$, which for 6 groups gives $P = \cdot435$, or a reasonable probability for the system as a whole. This test again confirms our result. Moreover, not one of the correlation ratios shown in this table, either of manual on visual characters or *vice versa*, is significant when compared with the difference of $\eta^2 - \overline{\eta}^2$ and the probable error of $\overline{\eta}^2$. We may also consider the question of the independence of the lateralities of ocular and manual characters by testing the system of η^2's as a whole. Suppose we take

$$Z = \frac{\text{Observed } \eta^2 - \overline{\eta}^2}{\sigma_{\overline{\eta}^2}},$$

where $\overline{\eta}^2$ is the value to be expected on the basis of independence and $\sigma_{\overline{\eta}^2}$ is the standard deviation of $\overline{\eta}^2$, or of samples from a population showing independence. Then the values of Z should form approximately a distribution about a mean of zero with a standard deviation of unity. The probable error of Z will be accordingly:

$$\frac{\cdot 67449}{\sqrt{n}}.$$

Taking the thirty cases of manual on ocular characters:

$$\overline{Z} = -\cdot 1168 \pm \cdot 1231,$$

and again the thirty cases of ocular on manual characters:

$$\overline{Z} = -\cdot 0524 \pm \cdot 1231.$$

Thus $\overline{Z}$ is in both cases negative or the weighted mean η^2 is below the weighted mean $\overline{\eta}^2$. In both cases also the value of $\overline{Z}$ differs from zero by less than its probable error. We may therefore conclude that the whole system of correlations between manual and ocular characters is consistent with absolute independence between manual and ocular lateralities. Thus all our investigations of the relationship of laterality of hand and eye entirely fail to confirm the theory, held by many authors, of universal absolute laterality, that is a superior eye determining which shall be the superior hand.

From our study of five different ocular and six manual characters we can safely say that there is no association either in coincidence or in intensity between the dominance of any character in the eye and any character in the hand which we have studied.

Problem 5. *Is either Manual or Ocular Dominance in itself a clear cut characteristic? Is it possible to assert for every character of both hand and eye that an individual is either dextral or sinistral? Is ambilaterality non-existent, or does it occupy an appropriate position midway between the easily recognised cases of dextrality and sinistrality?*

The answer to the questions involved in this problem are in a certain sense matters of definition. If we measure any anthropometric character of hand or eye, we find a continuous distribution of values forming what is termed a frequency curve. If we take the difference of characters on the right and left sides of the body, whether they be manual or ocular, we also obtain a frequency curve. In the characters we have dealt with in this paper, these frequency curves of laterality are not bi-modal; we have a continuous curve rising from extreme dextrality through a modal value, which may or may not approach the zero value of the variate, the curve afterwards falling again to extreme sinistrality. In other words the *ordinate* of this curve at zero difference occupies its *proper* place in the continuous series of ordinates.

This is illustrated by Diagrams XIV and XV for laterality distribution in the cases of Laterality of Grip and Laterality of Visual Acuity.

Now if we defined ambilaterality to arise when the difference was exactly zero, the occurrence of such cases would be practically impossible, if our measuring

instruments were indefinitely finely graduated. In other words "ambilaterality" would be solely a dividing line between plus and minus differences. Accordingly, any individual would be certainly either a dextralist or sinistralist for every character. But such a definition is for practical purposes absolutely idle. With a like definition emmetropia becomes merely the border line between myopia and hypermetropia, or indeed myopia itself a border line between two classes of myopic astigmatism. Yet ophthalmologists make substantial groups of emmetropic and myopic individuals. It is perfectly true that these groups, as those of ambilaterality, must be largely conventional, depending on the differences we find it of practical convenience to disregard. The fact of the arbitrary or conventional size of the group of ambilateralists is in the present case of little importance, for no dividing

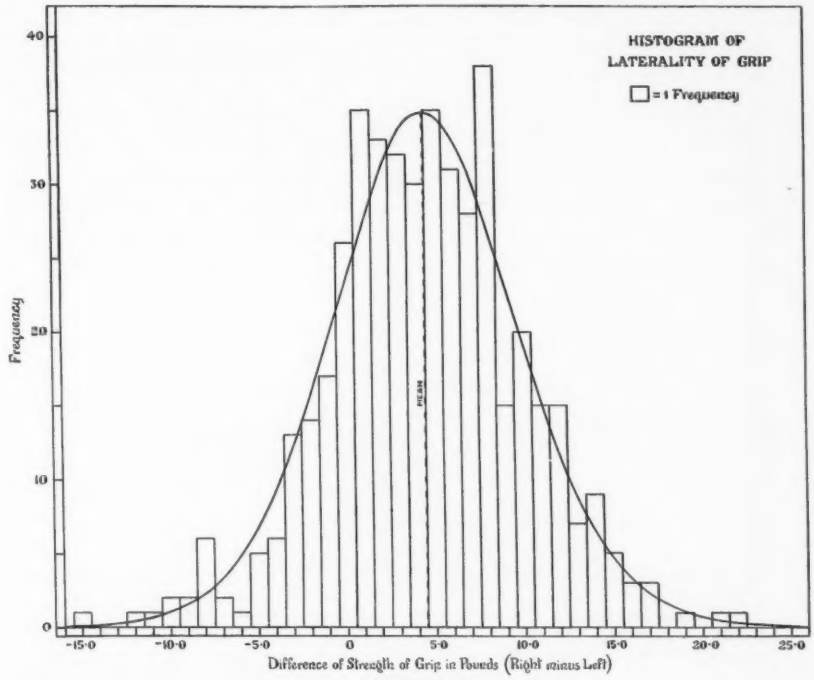


Diagram XIV.

in any proportions of those we have termed ambilateralists can possibly bring into anything like accordance all the various percentages found for dextralists and sinistralists. We may illustrate this by such a character as Visual Acuity, read by Snellen's types, converting the ratios into decimals. Here the range chosen for ambilaterality was that containing differences from $-\cdot015$ to $+\cdot015$ and the percentages were:

Dextralists: 39·14 °/₀; Ambilateralists: 16·47 °/₀; Sinistralists: 44·39 °/₀;

but had we taken our range of Ambilaterality from $-\cdot045$ to $+\cdot045$ we should have had:

Dextralists: 30·31 °/₀; Ambilateralists: 37·23 °/₀; Sinistralists: 32·46 °/₀.

The distinction between our standpoint and our observations and those of van Biervliet is *not* in the proportion of ambilateralists, but in the nature of the frequency curves for the differences. He asserts an *absolute absence* of frequency for a long range round zero difference; we find a continuous frequency curve covering the whole of this range round zero difference, and the number of persons to be treated as ambilateralists depends: (i) on the fineness with which we can measure our characters, or on the coarseness necessary or profitable in practice,

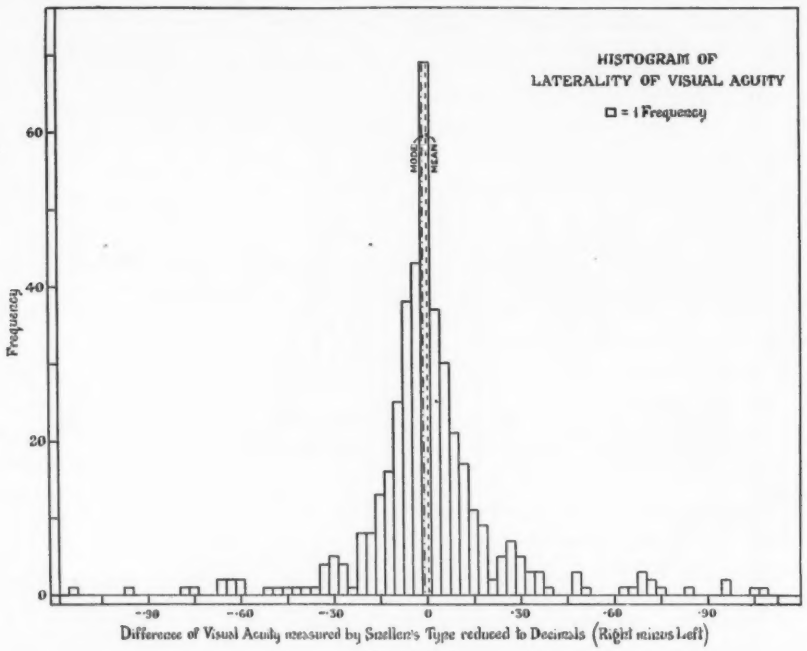


Diagram XV.

(ii) on the fact that no character in a human being remains absolutely constant; the pull of an ambilateralist may be greater on one trial with the left and on another with the right hand, and (iii) the personal equation or error of the recorder, e.g. it is no use trying to read refractions closer than $\frac{1}{4}$ dioptre; for if we did, our ambilateralists for refraction would be merely a function of the personal equation of the ophthalmologist. Hence in our record of the frequencies and corresponding percentages of ambilateralists, we have chosen ranges for grouping, which signify that there is *practically* no difference in the efficiency of right or left side; it would be idle to term a man a dextralist who could on the average pull half an ounce

more with his right than with his left hand, or a sinistralist who could decipher diamond type a tenth of an inch further off with his left than with his right eye.

The chief source of the different percentages found for the various characters arises, we believe, from an essential difference in the types of their frequency distributions; they may be symmetrical about zero difference, or skewed to the right or to the left. In other words if we had an adequate number of observations, we might determine the frequency curve for each character difference, and then integrate its areas from zero difference in both sinistral and dextral directions, we should thus reach absolute dextral and sinistral percentages only, with no ambilaterality. Short of this there is no reasonable method of truly proportioning the ambilateralists, who form the central group on a symmetrical subrange round zero. But such a proceeding could not equalise our proportional percentages for all ocular and manual characters, and what is more it would destroy the important and useful group of ambilaterality, the individuals with *practically* equal lateralities.

Problem 6. *To consider to what extent the Correlation of Lateralities in two Characteristics depends on the absolute Organic Correlations.*

The last question we have to consider is the connection which exists between the correlation of the intensities of laterality for two characters and the direct and cross organic correlations of those characters.

Let X = difference of right and left value for first character,

and Y = " " " " " " " second "

Let x_1 and y_1 be the absolute values of the two characters on the right side and x_2 and y_2 of the same characters on the left. Let σ_{x_1}, σ_{y_1}, σ_{x_2}, σ_{y_2} be the corresponding standard deviations. Then as usual we have:

$$r_{XY} = \frac{r_{x_1y_1}\sigma_{x_1}\sigma_{y_1} + r_{x_2y_2}\sigma_{x_2}\sigma_{y_2} - r_{x_1y_2}\sigma_{x_1}\sigma_{y_2} - r_{x_2y_1}\sigma_{x_2}\sigma_{y_1}}{\sqrt{\sigma_{x_1}^2 + \sigma_{x_2}^2 - 2r_{x_1x_2}\sigma_{x_1}\sigma_{x_2}}\sqrt{\sigma_{y_1}^2 + \sigma_{y_2}^2 - 2r_{y_1y_2}\sigma_{y_1}\sigma_{y_2}}} \quad \ldots\ldots\ldots(\epsilon_1).$$

Thus r_{XY} is known if we ascertain

(i) the variabilities of the absolute measurements,

(ii) the organic correlations $r_{x_1y_1}$ and $r_{x_2y_2}$,

(iii) the cross organic correlations $r_{x_1y_2}$ and $r_{x_2y_1}$,

(iv) the bilateral organic correlations $r_{x_1x_2}$ and $r_{y_1y_2}$.

Now as a first approximation we might assume the variabilities on right and left sides to be roughly equal, i.e.

$$\sigma_{x_1} = \sigma_{x_2} \text{ and } \sigma_{y_1} = \sigma_{y_2}.$$

Hence: $$r_{XY} = \frac{r_{x_1y_1} + r_{x_2y_2} - r_{x_1y_2} - r_{x_2y_1}}{2\sqrt{1 - r_{x_1x_2}}\sqrt{1 - r_{y_1y_2}}}, \text{ approximately} \ldots\ldots\ldots\ldots(\epsilon_2).$$

But there is some reason for believing that for various organs

$$r_{x_1y_2} = \tfrac{1}{2}(r_{x_1y_1}r_{y_1y_2} + r_{x_2y_2}r_{x_1x_2}), \quad r_{x_2y_1} = \tfrac{1}{2}(r_{x_2y_2}r_{y_1y_2} + r_{x_1y_1}r_{x_1x_2}) \ldots\ldots(\epsilon_3),$$

approximately.

Accordingly:

$$r_{XY} = \frac{r_{x_1y_1}(1-\frac{1}{2}r_{y_1y_2}-\frac{1}{2}r_{x_1x_2})+r_{x_2y_2}(1-\frac{1}{2}r_{x_1x_2}-\frac{1}{2}r_{y_1y_2})}{2\sqrt{1-r_{x_1x_2}}\sqrt{1-r_{y_1y_2}}}$$

$$= \tfrac{1}{2}(r_{x_1y_1}+r_{x_2y_2})\frac{(1-\frac{1}{2}r_{x_1x_2}-\frac{1}{2}r_{y_1y_2})}{\sqrt{(1-r_{x_1x_2})(1-r_{y_1y_2})}}, \text{ approximately.}$$

Further, suppose that the correlations on the two sides of the body are the same, i.e. $r_{x_1y_1}=r_{x_2y_2}=r_{xy}$, say, and the correlations of homologous organs are all the same, i.e. $r_{x_1x_2}=r_{y_1y_2}$. These things are not far from the truth. For example: Right and Left Femora = ·976 and Right and Left Tibiae = ·963, middle phalanges of middle and ring fingers on right hand = ·937 and on left hand = ·963. Also as a rule homologous organs have a correlation of about ·9. Consequently we should have:

$$r_{XY} = r_{xy}\frac{(1-2\times\frac{1}{2}\cdot 9)}{\sqrt{(1-\cdot 9)^2}} = r_{xy},$$

or the correlation of differences would equal the correlation of absolute characters. And thus, if, as is probably true, there was a very small or zero correlation between ocular and manual absolute characters, we should find no correlation, as we do find, between ocular and manual lateralities.

It seems worth while inquiring where the above relationship fails in the case of one or two instances. Is it because right and left side variabilities are not approximately equal? Or, is it because the formula for cross organic correlation is not adequately approximate?

The data for three cases are given in Table XXIII on p. 145. From these it appears that the right hand is significantly more variable than the left for Grip and the left hand significantly more variable than the right for Steadiness. There is no significant difference between right and left side variability for First Pull, Balancing or Near Point. Accordingly only in the third case, and then only neglecting differences of the order of the probable errors of random sampling, would it be legitimate to use a formula like (ϵ_3) instead of (ϵ_1).

Formula (ϵ_3) in the case of Grip and Initial Pull gives $r_{x_1y_2}=$ ·4213 and $r_{x_2y_1}=$ ·4204 as against the observed values ·4698 and ·5435 respectively. The former might just pass, the latter is too divergent for us to admit (ϵ_3) in this case. In the case of Steadiness of Hand and Balancing we have ·1299 and ·1312 as against ·2091 and ·2769 for $r_{x_1y_2}$ and $r_{x_2y_1}$ respectively as observed.

Lastly for Near Point and Initial Pull we have ·0239 and ·0233 as against ·0634 and ·0112 respectively. It is fairly clear therefore that we cannot make use of (ϵ_3) to simplify (ϵ_1). The last column of Table XXIII shows that (ϵ_1) gives results in reasonable accordance with the direct observation and application of the product moment method, but as it involves finding six correlation coefficients instead of one, it is hardly likely to be used unless these coefficients are needed for other purposes. It may be noted that the cross correlations $r_{x_1y_2}$ and $r_{x_2y_1}$ come far more nearly equal by (ϵ_3), than if they are found by direct observation.

TABLE XXIII.

Comparison of the Correlations of Lateralities obtained directly and from the Absolute Correlations.

Characters	Standard Deviations				Correlation Coefficients						
	σ_{x_1} for right side of first character	σ_{x_2} for left side of first character	σ_{y_1} for right side of second character	σ_{y_2} for left side of second character	$r_{x_1 x_2}$	$r_{y_1 y_2}$	$r_{x_1 y_1}$	$r_{x_2 y_2}$	$r_{x_1 y_2}$	$r_{x_2 y_1}$	$r_{(x_1-x_2)(y_1-y_2)}$
Grip and First Pull	8·1895lbs. ± ·1953	7·3052 lbs. ± ·1742	7·2253 lbs. ± ·1723	7·3608 lbs. ± ·1756	·7404 ± ·0141	·8096 ± ·0109	·5560 ± ·0218	·5300 ± ·0228	·4698 ± ·0247	·5435 ± ·0223	·1686 ± ·0308* ·1728 ± ·0308†
Steadiness of Hand and Balancing	2·5527‡ ± ·0609	2·8083 ‡ ± ·0670	7·1948 ‡ ± ·1716	7·2371 ‡ ± ·1726	·5125 ± ·0231	·5531 ± ·0218	·2132 ± ·0301	·2768 ± ·0291	·2091 ± ·0302	·2769 ± ·0292	·0040 ± ·0315* ·0046 ± ·0315†
Near Point and First Pull	1·1652 ± ·0278 inches	1·1519 ± ·0275 inches	7·0979 ± ·1693 lbs.	7·2746 ± ·1735 lbs.	·8282 ± ·0104	·7404 ± ·0141	·0225 ± ·0333	·0377 ± ·0335	·0634 ± ·0334	·0112 ± ·0331	− ·0354 ± ·0335* − ·0361 ± ·0335†

* Values found directly.

† Coefficients calculated from Formula (ϵ_1).

‡ Units those of marking scores.

IV. Summary and Conclusions.

From the results we have discussed above in detail the following general conclusions may be drawn:

(1) On the basis of our present data it is not true that all characters of the hand and eye are universally superior on one side and that this dominance is always of the same definite amount, namely in the ratio of 10 to 9, as van Biervliet held in his researches. The superiority of Grip in favour of the right hand is large to the average amount of 4·5 lbs., and for Steadiness and Balancing the results are about 1·5 times better in the right hand. But for other characters the degree of superiority of right side over the left is comparatively small, only measured in the decimal place of the unit adopted, or it may even be definitely insignificant.

(2) There is a large variability in the percentage occurrence of each type of laterality for the various characters tested. It has been shown that none of the percentages of dextrality agrees closely with the three-quarters dominance of the Mendelian ratio, and similarly there is no single percentage corresponding to the recessive Mendelian quarter except in the case of Horizontal Refraction which quite possibly arises from the fluctuations of random sampling. The frequencies of ambilaterality occurring for every character are so large that none of them can be regarded as non-existent if its probable error be taken into consideration.

(3) As far as our data are concerned it is impossible to say that persons dominant on one side for one manual or for one ocular character will be dominant on the same side for another of the same group, and still more impossible to say that they will be dominant on the same side for both groups. The mean percentage of same laterality, even if we include the frequency of the same ambilaterality either for the manual or the visual group, fails to exceed half of the whole population. That is to say, on the average, more than half the individuals of the total population will be dextral in one character but sinistral or ambilateral in others.

(4) The interrelations of the dominance of the five ocular characters are very variable. The lateralities of Visual Acuity and Refraction, taken either for the vertical or the horizontal meridian, are associated to a moderate extent (both = ·206) and are both significant in comparison with their probable errors. The association between the laterality of the two Refractions is ·66, a substantial relationship. There are three other correlations which just tend to be significant, two for the cases of the dominance of Near Point in relation with that of Visual Acuity and the Horizontal Refraction, and another for the dominance of Vertical Refraction and General Astigmatism. The other four correlation coefficients for the remaining pairs of visual characters are too small to be significant.

(5) The interrelations of the dominance of six manual characters are also very variable. We find two high associations with opposite sign in the manual group, one is positive in the case of the dominance of First Pull and Lack of Endurance (+ ·606) and the other negative in the case of the dominance of Last Pull and Lack of Endurance (− ·636). There are significant and positive correlations of ·17 and ·24

between the dominance of First Pull with those of Grip and of Last Pull. The other eleven correlations are probably not significant and some are lower than the values of their probable errors.

Not only is there a high association between the dominance of First Pull and Lack of Endurance in pulling but also two large correlations have been found for the absolute measures of these two variates for both hands. It seems probable that the correlations between Dominance of Lack of Endurance and Dominance of First and of Last Pull are due to spurious correlation as the sign indicates.

(6) The percentage occurrence of each type of laterality compared with its corresponding expected value on the theory of no association for any two characters, whether in the manual group or in the visual one, shows a fair correspondence with the relationship as dealt with by the correlation method.

(7) There is little truth in the theory of absolute laterality. It has been found that among thirty samples of one ocular character related with one manual one, no single value of the associational constants, whether linear or skew, can be treated as significant. The observed mean r is $\cdot0012 \pm \cdot0045$, which is not significantly different from the expected mean r, zero, and this is confirmed by different methods of testing.

The same is true in the comparison of the percentage occurrence for any pair of characters of hand and eye. It has been found that the difference between the observed percentages and the expected ones on the basis of no association is in general insignificantly small, viz. the frequency of laterality in hand and eye is quite independent.

(8) By considering the frequency of laterality for any two characters of hand and eye, the absolute theory can also be disproved. The mean percentage of the same laterality shown in the table is $33\cdot02 \pm \cdot72$ per cent. against $26\cdot57 \pm \cdot90$ per cent. of average opposite laterality and $40\cdot41 \pm 1\cdot45$ per cent. of average mixed laterality. In other words, two-thirds of the general population practically fail to follow the absolute theory.

(9) The classification and frequency of the three lateralities are not clear-cut characteristics varying with the character treated. They may be affected to a certain extent by such essential factors as the magnitude of the unit of measurement and the form of the frequency for a given trait. No particular person can be designated definitely as a dextralist, sinistralist or ambilateralist. He may change his type of laterality from one to another character under different experimental tests. The frequency of ambilaterality has really a magnitude appropriate to the value of the zero difference; it lies in a continuous distribution with a positive varying frequency of dextrality on one side and a negative one of sinistrality on the other.

(10) The intensity of the correlation of two absolute characters is not in itself adequate to determine the degree of laterality of the same characters. The intensity of the latter correlation depends on the values of four standard deviations

and three pairs of correlation coefficients for each side as described above. Therefore, the correlation of laterality cannot be predicted merely from the correlation of absolute measures.

Our data including eleven fundamental characters of hand and eye with an adequate number of cases seem much wider and more reliable than those with which van Biervliet and Parson have dealt. We believe that the present material can thus form a sufficient basis for solving the problems which have been previously raised. But all the results shown are as opposed to the theory of universal laterality as we had found them to be in Galton's data. We venture to assert that that theory, tested in many ways in the present paper and shown to fail, cannot be accepted on the ground of more sparse data treated by less adequate statistical methods in earlier memoirs. It is at any rate far from proven at present.

In conclusion I should like to express my great indebtedness to the editor, Professor Karl Pearson, for his helpful suggestions and criticisms in connection with this paper, and to Miss Ida McLearn for her careful preparation of the diagrams.

ON A METHOD OF ASCERTAINING LIMITS TO THE ACTUAL NUMBER OF MARKED MEMBERS IN A POPULATION OF GIVEN SIZE FROM A SAMPLE.

By KARL PEARSON.

THE problem suggested in the above title is a very fundamental one in statistical practice, but I have not seen the current method of dealing with it critically examined or justified. Suppose a population to number N individuals, of whom p are actually marked by a special characteristic and q not so; thus $N=p+q$. Now suppose a sample n taken of this population, and that in this sample r are found marked and s not so. Then the usual method is to consider the percentage of marked individuals to be

$$100\frac{r}{n} \pm 67{\cdot}449\sqrt{\frac{rs}{n^3}} \quad \text{.....(i),}$$

and to suppose that the probable error thus determined will measure in a rough sort of way the possible deviation of the sample value $100r/n$, from the actual, and unknown $100p/N$.

I do not think the reasoning leading to this result is satisfactory; first, because the result is independent of the size of the population sampled, and secondly because it really throws us back on the normal curve as representing the binomial, and this will only be correct if r or s be not small as compared with n.

The problem before us may really be said to consist practically of two somewhat different propositions:

(1) We draw a sample $n=r+s$, and ask on the basis of this experience, what will be the distribution of r' and s' in further samples of size $n'=r'+s'$.

This problem involves an appeal to Bayes' Theorem. I discussed it at some length in 1907*, in a paper entitled: "On the Influence of Past Experience on Future Expectation," and will only refer here to that paper.

(2) We draw a sample $n=r+s$, and ask on the basis of this experience what is the likelihood of various values of p and q in the actual population N. In other words, knowing r and s, we seek the distribution of p and q.

This is the problem I want to discuss on the present occasion, and it has to be clearly distinguished from (1).

* *Philosophical Magazine*, March, 1907, pp. 365—378.

If the actual numbers in the sampled population were p and q, then the chance of drawing r and s in a sample of n, out of N, is

$$C_{r,p} = \frac{\lfloor n}{\lfloor r \lfloor s} \frac{p(p-1)\dots(p-r+1)\,q(q-1)\dots(q-s+1)}{N(N-1)\dots\dots(N-n+1)}$$

$$= \frac{\lfloor n}{\lfloor r \lfloor s} \frac{\lfloor p}{\lfloor p-r} \frac{\lfloor q}{\lfloor q-s} \frac{\lfloor N-n}{\lfloor N} \quad \dots\dots\dots\dots\text{(ii)}.$$

Now the possible values of p are from $p=r$ to $p=N-s$, corresponding to $q=N-r$ to $q=s$. Hence, by the theory of inverse probabilities*, the probability of p given r in a sample, i.e. $C_{p,r}$, say, will be proportional to the above expression, or writing the constant as C_0, will be

$$C_{p,r} = C_0 \frac{\lfloor p}{\lfloor p-r} \frac{\lfloor q}{\lfloor q-s} \quad \dots\dots\dots\dots\dots\dots\dots\dots\text{(iii)}.$$

The series accordingly representing the various chances that the actual population contains r, $r+1$, $r+2 \dots N-s$ marked individuals is given by

$$S_{p,r} = C_0 \left\{ \frac{\lfloor r}{\lfloor 0} \frac{\lfloor N-r}{\lfloor N-n} + \frac{\lfloor r+1}{\lfloor 1} \frac{\lfloor N-r-1}{\lfloor N-n-1} + \frac{\lfloor r+2}{\lfloor 2} \frac{\lfloor N-r-2}{\lfloor N-n-2} + \dots \right\}$$

$$= \frac{C_0 \lfloor r}{\lfloor N-n} \frac{\lfloor N-r}{} \left\{ 1 + \frac{r+1}{\lfloor 1} \frac{N-n}{N-r} + \frac{(r+1)(r+2)}{\lfloor 2} \frac{(N-n)(N-n-1)}{(N-r)(N-r-1)} \right.$$

$$\left. + \dots + \frac{\lfloor s}{\lfloor r} \frac{\lfloor N-s}{\lfloor N-r} \right\} \quad \dots\dots\dots\dots\text{(iv)},$$

the series having a finite number of terms.

Disregarding the constant, we have clearly to deal with a hypergeometrical series, wherein

$$\left.\begin{array}{c} \alpha = r+1, \quad \beta = -(N-n), \quad \gamma = -(N-r) \\ \text{and} \qquad \gamma - \alpha - \beta - 1 = -(n+2) = \lambda, \text{ say} \end{array}\right\} \quad \dots\dots\dots\dots\text{(v)}.$$

We will now determine C_0 so that $S_{p,r}$ is equal to unity. This involves determining the sum of the finite hypergeometrical series in (iv). We can achieve this as follows: If a bag contain pM white and qM black balls and we take m balls out of the bag the distribution of 0, 1, 2 ... m black balls is given by the terms of the series

$$\frac{pM(pM-1)(pM-2)\dots(pM-m+1)}{M(M-1)(M-2)\dots(M-m+1)}$$

$$\times \left\{ 1 + \frac{m}{1} \frac{qM}{pM-m+1} + \frac{m(m-1)}{1.2} \frac{qM(qM-1)}{(pM-m+1)(pM-m+2)} \right.$$

$$\left. + \frac{m(m-1)(m-2)}{1.2.3} \frac{qM(qM-1)(qM-2)}{(pM-m+1)(pM-m+2)(pM-m+3)} + \dots \right\},$$

and the total sum must of course be unity. Hence we have

$$\frac{\lfloor M}{\lfloor M-m} \frac{\lfloor pM-m}{\lfloor pM} = 1 + \frac{m}{1} \frac{qM}{pM-m+1} + \frac{m(m-1)}{1.2} \frac{qM(qM-1)}{(pM-m+1)(pM-m+2)} + \dots.$$

* See Appendix I to this paper.

Put $\alpha=-m$, $\beta=-qM$, $\gamma=pM-m+1$ and we have, replacing factorials by Γ-functions,

$$\frac{\Gamma(\gamma-\alpha-\beta)\,\Gamma(\gamma)}{\Gamma(\gamma-\alpha)\,\Gamma(\gamma-\beta)}=1+\frac{\alpha\beta}{1!\,\gamma}+\frac{\alpha(\alpha+1)\beta(\beta+1)}{2!\,\gamma(\gamma+1)}+\ldots.$$

This is the sum of the series on the right-hand side supposing that series to be finite or converging.

Now in our special case

$$\alpha=r+1,\quad \beta=-(N-n),\quad \gamma=-(N-r),$$

and all the Γ-functions on the left have negative arguments. They must therefore be raised by $\Gamma(n)=\Gamma(n+1)/n$ to positive Γ-functions. We have

$$\Gamma(n)=\frac{1}{n}\frac{1}{n+1}\ldots\frac{1}{n+p-1}\Gamma(n+p).$$

Now let $n=-p+\epsilon$.

$$\Gamma(-p+\epsilon)=\frac{1}{-p+\epsilon}\,\frac{1}{-p+\epsilon+1}\ldots\frac{1}{+\epsilon-1}\Gamma(\epsilon)$$

$$=\frac{(-1)^p\,\Gamma(\epsilon)}{(p-\epsilon)(p-\epsilon-1)\ldots(1-\epsilon)}.$$

Now in our case

$\alpha=r+1$, $\beta=-(N-n)$, $\gamma=-(N-r)$, $\gamma-\alpha-\beta-1=-(n+2)=\lambda$, say,

$$\gamma-\alpha-\beta=-(n+1),\quad \gamma-\alpha=-(N+1),\quad \gamma-\beta=-(n-r)\quad\ldots\ldots\text{(vi)}.$$

Hence

$$\frac{\Gamma(\gamma-\alpha-\beta+\epsilon)\,\Gamma(\gamma+\epsilon)}{\Gamma(\gamma-\alpha+\epsilon)\,\Gamma(\gamma-\beta+\epsilon)}$$

$$=\frac{(-1)^{n+1}\,\Gamma(\epsilon)\times(-1)^{N-r}\,\Gamma(\epsilon)}{(n+1-\epsilon)(n-\epsilon)\ldots(1-\epsilon)(N-r-\epsilon)(N-r-\epsilon-1)\ldots(1-\epsilon)}$$

$$\frac{(-1)^{N+1}\,\Gamma(\epsilon)\times(-1)^{n-r}\,\Gamma(\epsilon)}{(N+1-\epsilon)(N-\epsilon)\ldots(1-\epsilon)(n-r-\epsilon)(n-r-\epsilon-1)\ldots(1-\epsilon)}$$

$$=\frac{(N+1-\epsilon)(N-\epsilon)\ldots(1-\epsilon)\times(n-r-\epsilon)(n-r-\epsilon-1)\ldots(1-\epsilon)}{(n+1-\epsilon)(n-\epsilon)\ldots(1-\epsilon)\times(N-r-\epsilon)(N-r-\epsilon-1)\ldots(1-\epsilon)}.$$

We can now put $\epsilon=0$ and we have

$$\frac{\Gamma(\gamma-\alpha-\beta)\,\Gamma(\gamma)}{\Gamma(\gamma-\alpha)\,\Gamma(\gamma-\beta)}=\frac{\lfloor N+1}{\lfloor n+1}\,\frac{\lfloor n-r}{\lfloor N-r}$$

for the sum of the finite hypergeometrical series in (iv). Hence if we choose $S_{p,r}$ to be unity

$$1=C_0\,\frac{\lfloor r}{\lfloor N-n}\,\frac{\lfloor N-r}{\lfloor N+1}\,\frac{\lfloor N+1}{\lfloor n+1}\,\frac{\lfloor n-r}{\lfloor N-r}\quad\text{or}\quad C_0=\frac{\lfloor N-n}{\lfloor N+1}\,\frac{\lfloor n+1}{\lfloor r}\,\frac{1}{\lfloor n-r},$$

or (iv) may be written

$$S_{p,r}=\frac{\lfloor n+1}{\lfloor N+1}\,\frac{\lfloor N-r}{\lfloor n-r}\left\{1+\frac{r+1}{\lfloor 1}\,\frac{N-n}{N-r}+\frac{(r+1)(r+2)}{\lfloor 2}\,\frac{(N-n)(N-n-1)}{(N-r)(N-r-1)}\right.$$

$$\left.+\ldots+\frac{\lfloor s}{\lfloor r}\,\frac{\lfloor N-s}{\lfloor N-r}\right\}\quad\ldots\ldots\ldots(\text{iv})^{\text{bis}}.$$

Turning to (iii) we may write it, since $n-r=s$,

$$C_{p,r}=\frac{n+1}{N+1}\left\{\frac{\lfloor n}{\lfloor r\,\lfloor s}\,\frac{\lfloor p\,\lfloor q}{\lfloor p-r\,\lfloor q-s}\,\frac{\lfloor N-n}{\lfloor N}\right\},$$

or we see from (ii) that

$$C_{p,r}=\frac{n+1}{N+1}C_{r,p},$$

i.e. the probability of the original population being p is to the probability of drawing a sample with r out of a population p as $n+1$ is to $N+1$. Since

$$C_{pr}=C_{rp}/S_{p=r}^{p=N-s}C_{r,p},$$

it follows that $$S_{p=r}^{p=N-s}C_{r,p}=(N+1)/(n+1).$$

The proof by which the value of C_0 is deduced may strike the reader as artificial, and may raise doubts in his mind. Actual numerical tests will soon convince him, however, that

$$1+\frac{r+1}{1}\frac{N-n}{N-r}+\frac{(r+1)(r+2)}{1.2}\frac{(N-n)(N-n-1)}{(N-r)(N-r-1)}+\ldots+\frac{\lfloor s}{\lfloor r}\frac{\lfloor N-s}{\lfloor N-r}=\frac{\lfloor N+1}{\lfloor n+1}\frac{\lfloor n-r}{\lfloor N-r}.$$

For example, if $N=15$, $n=10$, $r=4$, we have for the series

$$1+\frac{25}{11}+\frac{30}{11}+\frac{70}{33}+\frac{35}{33}+\frac{3}{11}+0=\frac{104}{11},$$

but we have

$$\frac{\lfloor N+1}{\lfloor n+1}\frac{\lfloor n-r}{\lfloor N-r}=\frac{\lfloor 16}{\lfloor 11}\frac{\lfloor 6}{\lfloor 11}=\frac{104}{11}\text{ also.}$$

We can now express the moment coefficients of our series by throwing it by aid of (v) into the hypergeometrical form, for which these quantities are known. Thus:

$$F(\alpha,\beta,\gamma,1)=1+\frac{\alpha\beta}{1!\gamma}+\frac{\alpha(\alpha+1)\beta(\beta+1)}{2!\gamma(\gamma+1)}+\ldots\quad\ldots\ldots\ldots\text{(vii).}$$

Now if we write $\lambda=\gamma-\alpha-\beta-1$, then the distance α, μ_1' of the mean of the hypergeometrical series from its first term is

$$\mu_1'=\frac{\alpha\beta}{\lambda}\quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(viii),}$$

and the moment coefficients about the mean are given by*

$$\mu_2=\frac{\alpha\beta(\lambda+\alpha)(\lambda+\beta)}{\lambda^2(\lambda-1)}\quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(ix),}$$

$$\mu_3=\frac{\alpha\beta(\lambda+\alpha)(\lambda+\beta)(\lambda+2\alpha)(\lambda+2\beta)}{\lambda^3(\lambda-1)(\lambda-2)}\quad\ldots\ldots\ldots\ldots\ldots\ldots\text{(x),}$$

$$\mu_4=\frac{\alpha\beta(\lambda+\alpha)(\lambda+\beta)}{\lambda^4(\lambda-1)(\lambda-2)(\lambda-3)}\{\lambda^4+\lambda^3(3\alpha\beta+6(\alpha+\beta)+1)+\lambda^2(3\alpha\beta(\alpha+\beta)+6(\alpha+\beta)^2+6\alpha\beta)+\lambda(3\alpha^2\beta^2+18\alpha\beta(\alpha+\beta))+18\alpha^2\beta^2\}\quad\ldots\ldots\text{(xi).}$$

* See *Philosophical Magazine*, February, 1899, pp. 238—240.

Finally, the distance of the mode from the first term will be given by

$$\text{Mode} = \text{nearest integer to } \frac{(\alpha-1)(\beta-1)}{\gamma-\alpha-\beta+1} \quad \ldots\ldots\ldots\ldots\ldots\text{(xii)}.$$

Now it is extremely troublesome to compute the successive terms of a hypergeometrical series which may have several hundred or perhaps thousand terms, even if we begin with terms in the neighbourhood of the mean, and compute only those of practical significance. It is better therefore to replace the hypergeometrical by its appropriate frequency curve, and in calculating the probability of different constitutions for the sampled population, use the areas of this curve.

Applying (viii) to the above results we have

$$\text{Mean} = \frac{(r+1)(N-n)}{n+2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(xiii)}.$$

This result is of considerable interest. In the usual practice the mean would be taken $= rN/n$. This is only true approximately if r, N and n are large numbers and N very much larger than n.

$$\text{Mode} = \frac{r}{n}(N-n+1) \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(xiv)}.$$

Since $\frac{r}{n}$ is less than unity, the mode may be taken as $\frac{r}{n}(N-n)$, or the mode will be greater or less than the mean, according as the number of marked individuals in the sample is greater or less than 50 per cent. of the sample. We can easily see the influence of these formulae on the usual result by taking a numerical example. In a sample of 100, from a population of 1000, 20 are found with a marked character. We should then have: Mean as found from the ordinary theory $= 200$, mean as found from above theory $= 20 + 185{\cdot}3 = 205{\cdot}3$, and mode $= 20 + 180{\cdot}2 = 200{\cdot}2$. The most probable result of the sampling therefore is that the original population sampled contained 200 marked individuals, although deviations in excess of this would be more frequent than those in defect.

We now turn to the standard deviation. This is given by

$$\mu_2 = \sigma^2 = \frac{(r+1)(n-r+1)(N+2)(N-n)}{(n+2)^2(n+3)}$$

$$= \frac{r+1}{n+2}\,\frac{s+1}{n+2}(N+2)\left(\frac{N+3}{n+3}-1\right) \quad \ldots\ldots\ldots\ldots\text{(xv)}.$$

The ordinary theory would, I take it, suppose the squared standard deviation of the original population's frequency to be best given by

$$\sigma^2 = \frac{r}{n}\,\frac{s}{n}N \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(xv)}^{\text{bis}},$$

which differs very considerably from that of the present theory and does not provide for $\sigma = 0$, when $n = N$ or we take the whole population as our sample. It will be clear that when the sample is a small one, the variability of the distribution of possible actual populations is very large indeed.

In the case of the particular numerical example we have already dealt with, where $r = 20$, $n = 100$ and $N = 1000$, we find

$$\mu_2 = 1431{\cdot}4526, \text{ or } \sigma_p = 37{\cdot}8345.$$

On the old theory

$$\mu_2 = 160, \text{ or } \sigma_p = 12{\cdot}6491.$$

Naturally, however, σ_p will tend to become infinite with N, and what we really want is the variability of p/N or better in percentages $100\,p/N$. Thus the two ways of approaching the problem give:

	Mean Percentage	S.D. in °/₀
Present way	20·53	3·78
Usual way	20·00	1·26

which sufficiently emphasises that very different conclusions will be drawn from the two methods. Generally from (xv)

$$\sigma_{p/N} = \sqrt{\frac{(r+1)}{(n+2)}\frac{(s+1)}{(n+2)}\frac{1}{(n+3)}\left(1+\frac{2}{N}\right)\left(1-\frac{n}{N}\right)} \quad \text{.....(xvi).}$$

If N becomes indefinitely large, n remaining finite, this reduces to

$$\sigma_{p/N} = \sqrt{\frac{(r+1)(s+1)}{(n+2)^2(n+3)}} \quad \text{.....(xvi)}^{\text{bis}},$$

as against the usual formula

$$\sigma_{p/N} = \sqrt{\frac{rs}{n^3}}.$$

The deviation of this from (xvi)$^{\text{bis}}$ will be considerable, when we are dealing with *small* samples.

We now turn to the third moment coefficient in (x) and substitute the values from (v). We find

$$\mu_3 = \frac{(r+1)(s+1)(s-r)(N-n)(N+2)(N-n+N+2)}{(n+2)^3(n+3)(n+4)} \quad \text{...(xvii).}$$

Clearly μ_3 will be negative when $r > s$, i.e. greater than 50 per cent. of the sample, but this was the condition already found for the mode being greater than the mean or the skewness negative.

If M_3 be the μ_3 for p/N we must divide by N^3 or

$$M_3 = \frac{(r+1)(s+1)(s-r)}{(n+2)^3(n+3)(n+4)}\left(1-\frac{n}{N}\right)\left(1+\frac{2}{N}\right)\left(2-\frac{n-2}{N}\right) \quad \text{...(xvii)}^{\text{bis}},$$

or, for sampling an indefinitely large population, whatever the size of the sample,

$$M_3 = 2\,\frac{(r+1)(s+1)(s-r)}{(n+2)^3(n+3)(n+4)} \quad \text{.....(xvii)}^{\text{ter}}.$$

However, for curve fitting, it is the constant $\beta_1 = \mu_3^2/\mu_2^3$ that we generally use and not μ_3. For the hypergeometrical series in general we have

$$\beta_1 = \frac{(\lambda+2\alpha)^2(\lambda+2\beta)^2(\lambda-1)}{\alpha\beta(\lambda+\alpha)(\lambda+\beta)(\lambda-2)^2} \quad \text{.....(xviii),}$$

and this for our particular case becomes

$$\beta_1 = \frac{(s-r)^2}{(r+1)(s+1)} \frac{n+3}{(n+4)^2} \left(4 + \frac{(n+2)^2}{(N+n)(N+2)}\right) \quad \text{.....(xix),}$$

a fairly compact form.

For the special numerical case previously used as an illustration

$$\beta_1 = \cdot 08084,$$

which indicates a certain amount of skewness in the distribution of possible p's, but for 20 per cent. of marked individuals, the hypergeometrical series is approaching symmetry. Turning to μ_4 we find that for the general hypergeometrical series it may be written

$$\mu_4 = \frac{\alpha\beta(\lambda+\alpha)(\lambda+\beta)}{\lambda^4(\lambda-1)(\lambda-2)(\lambda-3)} \{\lambda^4 + \lambda^3(3\alpha\beta + 6(\alpha+\beta) + 1) + \lambda^2(3\alpha\beta(\alpha+\beta) + 6(\alpha+\beta)^2 + 6\alpha\beta) + \lambda(3\alpha^2\beta^2 + 18\alpha\beta(\alpha+\beta)) + 18\alpha^2\beta^2\} \quad \text{.....(xx).}$$

Or, again,

$$\mu_4 = \frac{\alpha\beta(\lambda+\alpha)(\lambda+\beta)}{\lambda^2(\lambda-1)(\lambda-2)(\lambda-3)} \left\{\frac{\lambda(\lambda+1) + 6\alpha(\lambda+\alpha) + 6\beta(\lambda+\beta) + 3(\lambda+6)\alpha\beta(\lambda+\alpha)(\lambda+\beta)}{\lambda^2}\right\} \quad \text{.....(xx)}^{\text{bis}}.$$

Hence

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3(\lambda+6)(\lambda-1)}{(\lambda-2)(\lambda-3)} + \frac{\lambda^2(\lambda-1)}{(\lambda-2)(\lambda-3)} \left\{\frac{\lambda(\lambda+1)}{\alpha\beta(\lambda+\alpha)(\lambda+\beta)} + \frac{6}{\beta(\lambda+\beta)} + \frac{6}{\alpha(\lambda+\alpha)}\right\} \quad \text{.....(xxi).}$$

For our special hypergeometrical this becomes

$$\beta_2 = \frac{3(n-4)(n+3)}{(n+4)(n+5)} + \frac{(n+2)^2(n+3)}{(n+4)(n+5)} \left\{\left(6 + \frac{(n+1)(n+2)}{(N-n)(N+2)}\right) \frac{1}{(r+1)(s+1)} - \frac{6}{(N+2)(N-n)}\right\} \quad \text{.....(xxi)}^{\text{bis}}.$$

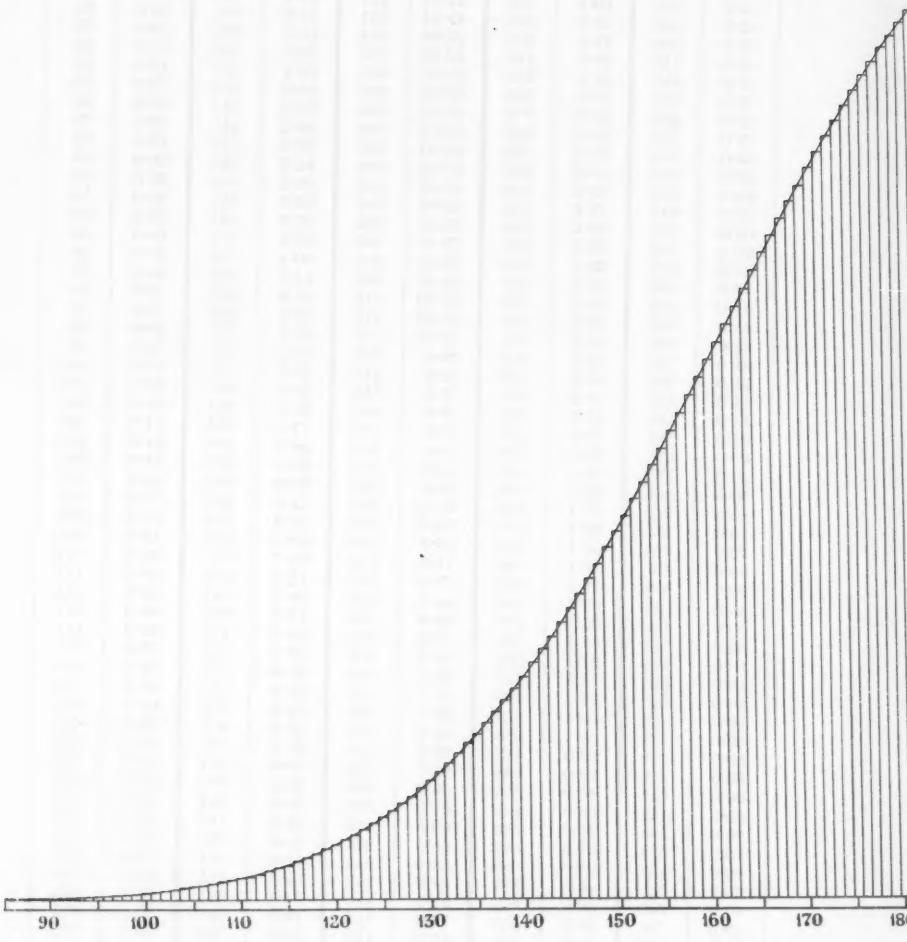
It seems desirable to apply this to a numerical example, and we will deal with that already discussed, in which we supposed a sample of 100, out of a population of 1000, showed 20 marked individuals and we required the probabilities of various numbers of marked individuals in the original population on the basis of this sample. We give first the frequencies of the various values of p*.

* The probabilities were found *in succession*, each from the preceding one; this required eight decimals; these were multiplied by 1000 and three figures cut off. The work is far more laborious than finding the curve in (xxiii), and plotting it. The planimeter will then give any required area with ample statistical accuracy.

Frequencies Per Mille of various possible Populations.

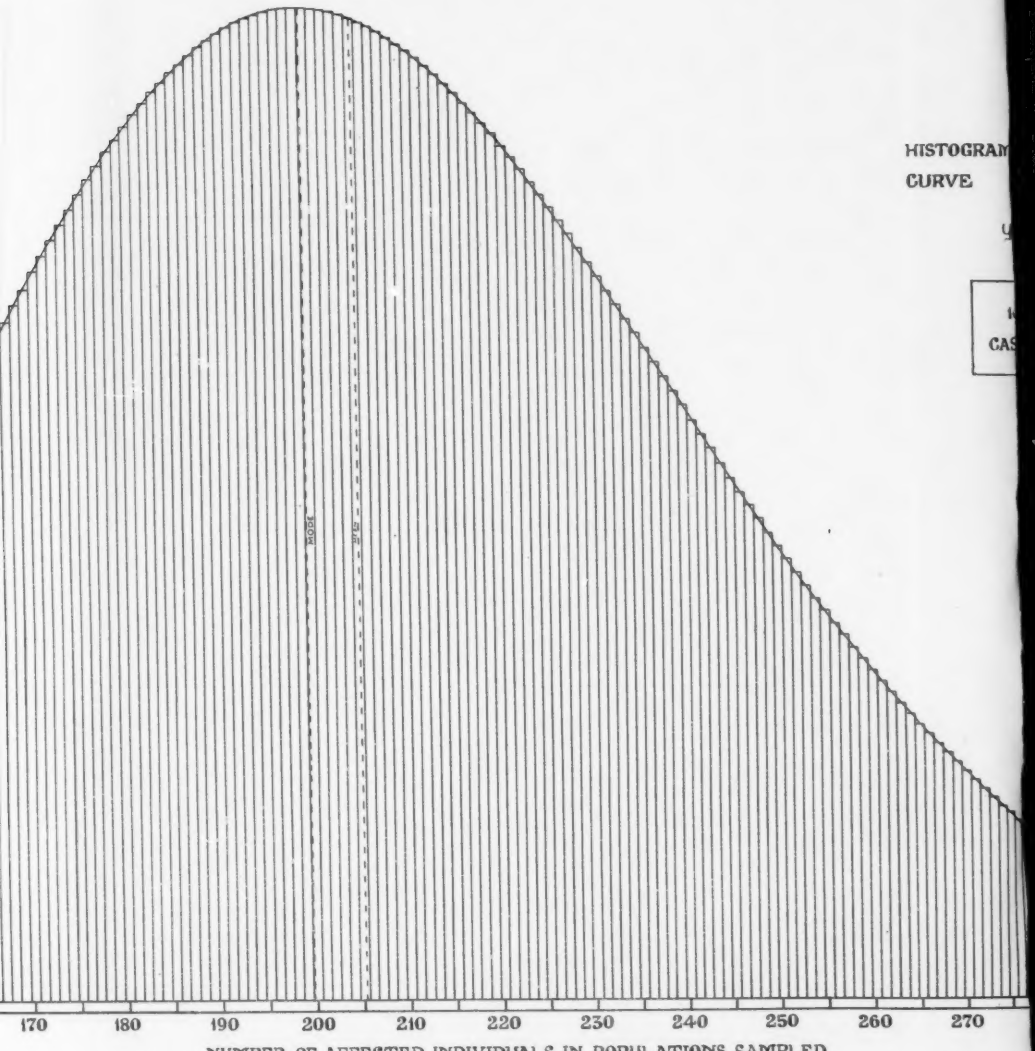
p	Per Mille	*p*	Per Mille	*p*	Per Mille	*p*	Per Mille	*p*	Per Mille
80	0·00	*137*	1·99	*194*	10·44	*251*	4·69	*308*	0·40
81	0·01	*138*	2·11	*195*	10·48	*252*	4·55	*309*	0·38
82	0·01	*139*	2·24	*196*	10·51	*253*	4·42	*310*	0·36
83	0·01	*140*	2·37	*197*	10·53	*254*	4·28	*311*	0·34
84	0·01	*141*	2·50	*198*	10·55	*255*	4·15	*312*	0·32
85	0·01	*142*	2·64	*199*	10·56	*256*	4·01	*313*	0·30
86	0·01	*143*	2·79	*200*	10·56	*257*	3·89	*314*	0·28
87	0·01	*144*	2·93	*201*	10·55	*258*	3·76	*315*	0·27
88	0·01	*145*	3·08	*202*	10·54	*259*	3·63	*316*	0·25
89	0·02	*146*	3·24	*203*	10·52	*260*	3·51	*317*	0·24
90	0·02	*147*	3·40	*204*	10·49	*261*	3·39	*318*	0·22
91	0·02	*148*	3·56	*205*	10·46	*262*	3·27	*319*	0·21
92	0·02	*149*	3·72	*206*	10·42	*263*	3·16	*320*	0·20
93	0·03	*150*	3·89	*207*	10·37	*264*	3·05	*321*	0·19
94	0·03	*151*	4·06	*208*	10·32	*265*	2·93	*322*	0·18
95	0·04	*152*	4·24	*209*	10·26	*266*	2·83	*323*	0·17
96	0·04	*153*	4·42	*210*	10·19	*267*	2·73	*324*	0·16
97	0·05	*154*	4·60	*211*	10·12	*268*	2·63	*325*	0·15
98	0·06	*155*	4·78	*212*	10·04	*269*	2·53	*326*	0·14
99	0·07	*156*	4·96	*213*	9·95	*270*	2·43	*327*	0·13
100	0·07	*157*	5·15	*214*	9·86	*271*	2·34	*328*	0·12
101	0·08	*158*	5·33	*215*	9·77	*272*	2·24	*329*	0·11
102	0·10	*159*	5·52	*216*	9·67	*273*	2·16	*330*	0·11
103	0·11	*160*	5·71	*217*	9·56	*274*	2·07	*331*	0·10
104	0·12	*161*	5·90	*218*	9·45	*275*	1·99	*332*	0·09
105	0·14	*162*	6·09	*219*	9·34	*276*	1·90	*333*	0·09
106	0·15	*163*	6·28	*220*	9·22	*277*	1·83	*334*	0·08
107	0·17	*164*	6·47	*221*	9·10	*278*	1·75	*335*	0·08
108	0·19	*165*	6·65	*222*	8·97	*279*	1·68	*336*	0·07
109	0·21	*166*	6·84	*223*	8·84	*280*	1·60	*337*	0·07
110	0·23	*167*	7·03	*224*	8·71	*281*	1·54	*338*	0·06
111	0·25	*168*	7·21	*225*	8·57	*282*	1·47	*339*	0·06
112	0·27	*169*	7·39	*226*	8·44	*283*	1·40	*340*	0·05
113	0·30	*170*	7·57	*227*	8·29	*284*	1·34	*341*	0·05
114	0·33	*171*	7·75	*228*	8·15	*285*	1·28	*342*	0·05
115	0·37	*172*	7·92	*229*	8·01	*286*	1·22	*343*	0·04
116	0·40	*173*	8·09	*230*	7·86	*287*	1·17	*344*	0·04
117	0·44	*174*	8·26	*231*	7·71	*288*	1·12	*345*	0·04
118	0·49	*175*	8·42	*232*	7·56	*289*	1·06	*346*	0·04
119	0·53	*176*	8·58	*233*	7·41	*290*	1·01	*347*	0·03
120	0·58	*177*	8·73	*234*	7·25	*291*	0·97	*348*	0·03
121	0·63	*178*	8·88	*235*	7·10	*292*	0·92	*349*	0·03
122	0·69	*179*	9·02	*236*	6·95	*293*	0·88	*350*	0·02
123	0·74	*180*	9·16	*237*	6·79	*294*	0·83	*351*	0·02
124	0·81	*181*	9·29	*238*	6·64	*295*	0·79	*352*	0·02
125	0·87	*182*	9·42	*239*	6·48	*296*	0·75	*353*	0·02
126	0·94	*183*	9·54	*240*	6·33	*297*	0·72	*354*	0·02
127	1·02	*184*	9·66	*241*	6·18	*298*	0·68	*355*	0·02
128	1·09	*185*	9·77	*242*	6·02	*299*	0·65	*356*	0·01
129	1·17	*186*	9·87	*243*	5·87	*300*	0·61	*357*	0·01
130	1·26	*187*	9·97	*244*	5·72	*301*	0·58	*358*	0·01
131	1·35	*188*	10·06	*245*	5·57	*302*	0·55	*359*	0·01
132	1·45	*189*	10·14	*246*	5·42	*303*	0·52	*360*	0·01
133	1·55	*190*	10·21	*247*	5·27	*304*	0·50	*361*	0·01
134	1·65	*191*	10·28	*248*	5·12	*305*	0·47	*362*	0·01
135	1·76	*192*	10·34	*249*	4·98	*306*	0·44	*363*	0·00
136	1·87	*193*	10·40	*250*	4·84	*307*	0·42		

SAMPLE OF 100 SHOWIN



To face p. 157

100 SHOWING 20 AFFECTED INDIVIDUALS. DISTRIBUTION OF POSSIBLE POPULATIONS OF 1000 S



HISTOGRAM
CURVE

NUMBER OF AFFECTED INDIVIDUALS IN POPULATIONS SAMPLED

Diagram I.

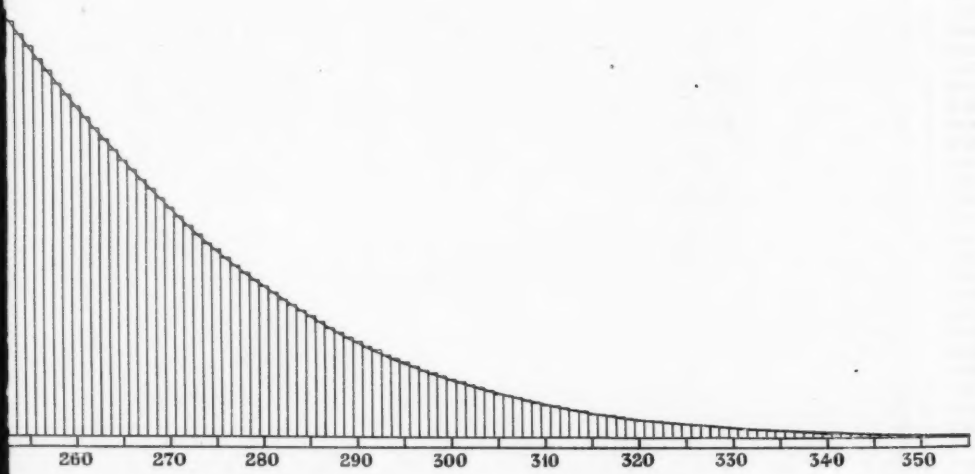
LATIONS OF 1000 SAMPLED.

HISTOGRAM = ACTUAL HYPERGEOMETRICAL SERIES

CURVE = TYPE I WITH EQUAL FIRST FOUR MOMENTS

$$y = 10{\cdot}5606\left(1+\frac{x}{189{\cdot}5250}\right)^{19{\cdot}9076}\left(1-\frac{x}{756{\cdot}9678}\right)^{79{\cdot}5115}$$

10 CASES	1000 POSSIBLE POPULATIONS



LD

Using the formulae (xv), (xix) and (xxi)^{bis} for σ^2, β_1 and β_2 respectively, we find the following Type I, limited range curve—where the distance from the origin (mode of curve) to the mean, taken as the same for curve and hypergeometrical, is 5·5950—is

$$y = 10{\cdot}5606\left(1+\frac{x}{189{\cdot}5250}\right)^{19{\cdot}9076}\left(1-\frac{x}{756{\cdot}9678}\right)^{79{\cdot}5116},$$

the origin of the curve being the mode at 199·6991. Diagram I shows the curve superposed on the hypergeometrical histogram; the accordance is remarkably good. But the labour of fitting the curve after we have found β_1 and β_2 is considerable. Examining the resulting curve we see that the two powers are singularly close to the values of r and s in the sample. Let us assume that they are such; and accordingly if b be the range of the curve, we shall have

$$a_1 = rb/n \text{ and } a_2 = sb/n,$$

and for the equation to the curve

$$y = y_0\left(\frac{rb}{n}+x\right)^r\left(\frac{sb}{n}-x\right)^s.$$

If M be the total number of possible populations from which the sample may have been drawn, usually taken as a thousand, then r and s are integers; we have

$$y_0 = \frac{M}{b^{n+1}}\frac{(n+1)!}{r!\,s!},$$

or the equation to the curve is

$$y = \frac{M}{b^{n+1}}\frac{(n+1)!}{r!\,s!}\left(\frac{rb}{n}+x\right)^r\left(\frac{sb}{n}-x\right)^s \quad \text{.....(xxii).}$$

It remains to consider a suitable method of determining the curve range b. The following suggest themselves:

(i) We might make the range of the curve equal to the range of the hypergeometrical. But the hypergeometrical is zero for $p = r-1$ and for $p = N-s+1$ or the total range is $N-n+2$. This is the value which should accordingly be given to b.

(ii) Our Diagram I indicates, however, that the practical range of both curve and hypergeometrical falls far short of the absolute range. We may therefore get a much better result by equating some physical constant of the curve and the series which is more important than the range for the useful part of the series. As a second possibility we might make the distance from mode to mean the same for curve and series, i.e.

$$\frac{b(s-r)}{n(n+2)} = \frac{(N-n)(s-r)}{n(n+2)}.$$

This gives $b = N-n$, which will usually be very close to the result in (i). Indeed, had we taken our range from first to last term in the hypergeometrical, the two ranges would be the same, and the distance from mode to mean the same for curve and series.

(iii) As (i) and (ii) are likely to succeed or fail together another mode of considering the matter is suggested. We may equate the standard deviations of curve and series. Now the standard deviation of the curve is $\frac{b^2(r+1)(s+1)}{(n+2)^2(n+3)}$, and of the series $\frac{(N+2)(N-n)(r+1)(s+1)}{(n+2)^2(n+3)}$. Hence we have

$$b^2 = (N+2)(N-n),$$

or

$$b = (N-n)\sqrt{1+\frac{n+2}{N-n}}.$$

This will accordingly give a somewhat larger range to the curve than in (i) or (ii), but has the advantage of making the variabilities in the curve and the series the same.

(iv) Still a further method was tried, namely, not fixing the powers of the two factors of the equation at r and s, but so choosing them that both range and standard deviation were the same for curve and series. This, however, was found to be more laborious, and to give less satisfactory results, than adopting (xxii) and choosing b appropriately. Finally, comparing the results obtained by choosing b as in (i) from the range or as in (iii) from the standard deviation, we found that for a series of curves the latter gave markedly the better result, and that the description by means of*

$$\left.\begin{aligned} y &= \frac{M}{b^{n+1}}\frac{(n+1)!}{r!\,s!}\left(\frac{rb}{n}+x\right)^r\left(\frac{sb}{n}-x\right)^s \\ \text{where}\quad b &= \sqrt{(N+2)(N-n)} \end{aligned}\right\} \quad \text{.....(xxiii),}$$

was really as good practically as if we had used the long formula for β_1 and β_2, and actually fitted the curve by its first four moments to the hypergeometrical series. We therefore propose to adopt (xxiii) as the practical solution of the problem of describing the distribution of possible sampled populations.

We will illustrate this by three examples.

Illustration I. Population of 1000. *Sample of* 100 *shows* 20 *affected individuals.*

Equation (xxiii) provides us with the following result:

$$y = y_0(189{\cdot}9263+x)^{20}(759{\cdot}7052-x)^{80},$$

where

$$\log y_0 = \overline{275}{\cdot}000{,}407.$$

Here the mean and standard deviation of series and curve are identical, and it takes a tenth of the labour to write down the equation to the curve that the fitting by moments takes. *It was not, however, possible to draw this curve on*

* For some purposes, for example, in determining the modal ordinate, it is convenient to write the equation to the curve in the form

$$y = \frac{M}{b}\frac{(n+1)!}{n^n}\frac{r^r}{r!}\frac{s^s}{s!}\left(1+\frac{n}{r}\frac{x}{b}\right)^r\left(1-\frac{n}{s}\frac{x}{b}\right)^s.$$

Thus the modal ordinate

$$y_{\text{mo.}} = \frac{M}{\sqrt{(N+2)(N-n)}}\frac{(n+1)!}{n^n}\frac{r^r}{r!}\frac{s^s}{s!},$$

and can be easily calculated logarithmically from tables of the log. factorials.

Diagram I, for it could not be shown as distinct from the moment curve on the scale of the diagram. We note the following results:

	Mode of Populations	Range
Actual Series	200·20	900
Curve fitted by Four Moments	199·70	946·5
Curve fitted by assuming powers of factors to be r and s	199·71	949·6

It will be seen that the mode is slightly out of position for both curves, but as the modal number can only be a *whole* number for all populations, it is 200 for all three cases, i.e. the number in the sample increased up to the size of the population. The curves give good fits to the bulk of the practical part of the series by extending the range beyond its actual value, but in this extension they provide no practically sensible frequency.

Illustration II. Population of 100. *Sample of* 40 *showing* 10 *affected individuals.*

The two curves have for their equations

Curve from Four Moments:

$$y = 75{\cdot}5688\left(1+\frac{x}{18{\cdot}5102}\right)^{8{\cdot}8616}\left(1-\frac{x}{54{\cdot}6906}\right)^{26{\cdot}1827}.$$

Curve from Equation (xxiii):

$$y = {\cdot}8177\times 10^{-64}(19{\cdot}5576+x)^{10}(58{\cdot}6728-x)^{30}.$$

The histogram of the hypergeometrical series is shown in Diagram II, but the two curves so nearly coincide that they cannot both be represented on a diagram of this size. It is clear that (xxiii) gives a perfectly adequate representation of the hypergeometrical series*. We note the following results:

	Mode of Populations	Range
Actual Series	25·25	60
Curve fitted by Four Moments ...	24·74	73·2
Curve fitted by Equation (xxiii) ...	24·78	78·2

The mean of the series is 25·714. It is clear that the modal value to the nearest whole number—which is all that can occur—is 25 for the series and both curves. It is further evident that the goodness of fit for the curves which have the same mean and standard deviation as the series is attained by prolonging the range into regions, which, if theoretically impossible, still do not for practical purposes provide any sensible frequency.

* If we make the range and standard deviation the same as for the hypergeometrical series, but let the powers in the equation differ from r and s (Method (iv)), we find

$$y = {\cdot}377787\times 10^{-32}(14{\cdot}4327+x)^{5{\cdot}3627}(43{\cdot}5673-x)^{16{\cdot}9314},$$

which gives an excess of frequency at the tails and a deficiency at the mode. The curve does not give a bad fit, but it is inferior to either of the two considered above. The same was the case in Illustration I when Method (iv) was applied.

SAMPLE OF 40 SHOWING 10 AFFECTED INDIVIDUALS. DISTRIBUTION OF 1000 POSSIBLE SAMPLED POPULATIONS OF 100

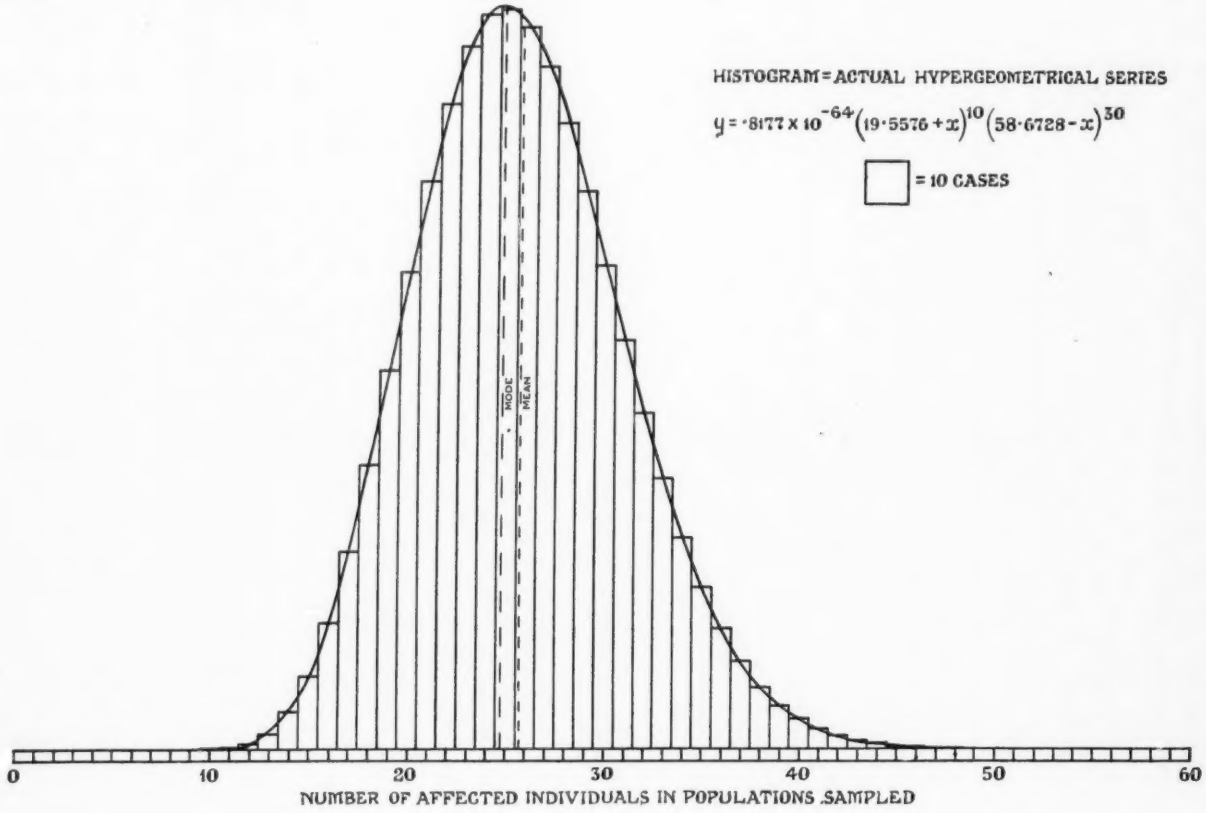


Diagram II.

Illustration III. Population of 30. *Sample of* 10 *showing* 2 *affected individuals.*

This illustration was taken to indicate what the power of description of the curves might be in still more extreme cases. The equations are:

Curve from Four Moments:

$$y = 131{\cdot}0512\left(1 + \frac{x}{4{\cdot}7158}\right)^{1{\cdot}7111}\left(1 - \frac{x}{19{\cdot}1170}\right)^{6{\cdot}9367}.$$

Curve from Equation (xxiii):

$$y = 131{\cdot}3094\left(1 + \frac{x}{5{\cdot}0596}\right)^{2}\left(1 - \frac{x}{20{\cdot}2386}\right)^{8}.$$

Diagram III shows both curves, which with small samples from small populations differ visibly in a diagram of this size. But the deviations of one from

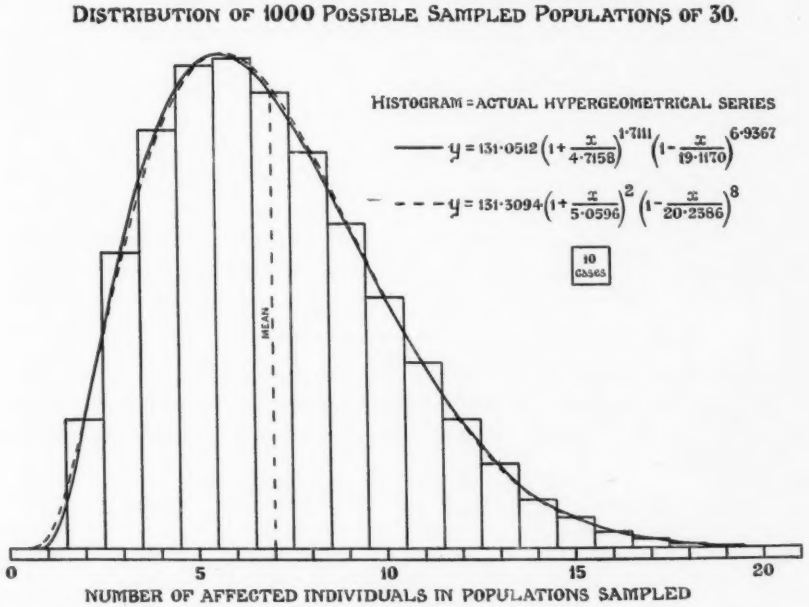

SAMPLE OF 10 SHOWING 2 AFFECTED INDIVIDUALS
DISTRIBUTION OF 1000 POSSIBLE SAMPLED POPULATIONS OF 30.
HISTOGRAM = ACTUAL HYPERGEOMETRICAL SERIES
y = 131·0512 (1 + x/4·7158)^1·7111 (1 − x/19·1170)^6·9367
y = 131·3094 (1 + x/5·0596)^2 (1 − x/20·2386)^8
10 CASES
MEAN
0
5
10
15
20
NUMBER OF AFFECTED INDIVIDUALS IN POPULATIONS SAMPLED

Diagram III.

the other and both from the histogram are not of a kind to be of any real importance in practical statistics. In this case the four-moment curve gives the somewhat better result.

The following Table giving the terms of the hypergeometrical and the ordinates of the two curves will enable the reader to judge the degree of accuracy with which even in an extreme case like the present the curves can be used to replace the series.

Distribution of 1000 *possible Populations of* 30, *when a Sample of* 10 *has shown two Marked Individuals.*

	Actual Hypergeometrical Series Value of Terms	Four Moment Curve Ordinates	Curve Equation (xxiii) Ordinates
Marked Individuals	—	[1·5 13·40]	[1·5 15·95]
2	36·71	34·68	34·89
3	78·66	78·65	76·41
4	110·70	111·27	109·47
5	127·74	128·25	127·61
6	130·29	130·02	130·87
7	121·60	121·28	122·38
8	105·74	105·34	106·48
9	86·52	86·23	87·00
10	66·95	66·84	67·16
11	49·09	49·26	48·98
12	34·11	34·25	33·66
13	22·39	22·55	21·81
14	13·83	13·96	13·05
15	7·98	8·05	7·85
16	4·26	4·28	4·19
17	2·07	2·06	2·04
18	·89	·88	·89
19	·33	·32	·34
20	·10	·09	·11
21	·02	—	—
22	·002	—	—
Total	999·98 [At Mode	5·6475 131·05	5·7350 131·31]

It is worth noting in this case that the hypergeometrical series has only a slight degree of leptokurtosis ($\beta_2 = 3{\cdot}07976$) but a considerable degree of skewness ($\beta_1 = {\cdot}37364$). We have the following results, the mean of the series and the two curves being 7.

	Mode of Populations	Range
Actual Series	6·2	20
Curve fitted by Four Moments ...	5·648	23·83
Curve fitted by Equation (xxiii) ...	5·735	25·30

Again we see that the mode as a whole number in the series and the curves is 6, and corresponds to the number of affected individuals in the sample, i.e. not the mean but the modal value in the sampled populations corresponds to the sample proportion. It will be further noticed that the goodness of fit of the curves to the series (not so good now at the start of the hypergeometrical series) is obtained by extending the range. This extension is of no importance practically on the right of the curves (see Diagram III), but it is of some importance on the left, where the curves instead of starting at 1·5 of the histogram of the hyper-

geometrical series start at 0·932 and 0·675 respectively. The small frequencies below 1·5 must be thrown into the frequency centred at 2 in such cases. It is possible that in this case our Method (iv) might give a closer result at this end of the series, but the general result is so satisfactory that it seems best to maintain a single method for all cases, namely, that of Equation (xxiii). This is so simple in application that it is preferable to the four-moment curve, although the latter gives a slightly better fit to the histogram.

Besides the importance of the present problem itself as providing a measure of the manner in which the sampled population may differ from the sample, our curve enables us readily to provide an asymptotic expression for any number of terms of a finite or of a converging hypergeometrical series, which may possibly be worthy of the consideration of the mathematician.

While attention has been frequently paid to the problem of future expectancy on the basis of past experience I do not know of anyone but Laplace who has attempted the present problem on the basis of inverse probabilities. In his treatment certain difficulties arise which I have considered in Appendix II to this paper.

I have to thank very cordially my colleagues Miss Margaret Moul and Miss Elizabeth Worthington for much arithmetical aid in this investigation of a simple curve to represent the hypergeometrical series. This aid is not fully represented in this paper, as several variations of Equation (xxiii) were tried from which the most satisfactory could only be picked after all the curves had been fitted to the series.

APPENDIX I.

NOTE ON THE THEORY OF INVERSE PROBABILITIES.

I am fully aware of the criticisms which have been raised on the use of the theory of inverse probabilities: see for example Venn's *Logic of Chance*, Chapter VII, §§ 9—18. But let us look into the actual needs of statistical science in ordinary practice. It is rare in anthropometry, especially in craniometry, in psychology or in medicine that a complete population is measurable; we must content ourselves with a sample. According to Dr Venn we cannot infer from a known sample anything as to the population sampled. We may draw conclusions for our sample but for nothing else. If we compare two samples of crania, one from Tibet and the other from Japan, we cannot assert anything about the difference between Tibetan and Japanese races as such, because we must not use the theory of inverse probabilities to determine some more or less approximate measure of the divergence of the populations from their samples. Or, may we be allowed to do so on the same ground as Dr Venn allows theory to be applied in the case of two samples of medical treatment to determine which in *general* is more efficient?

That is because:

The problem discusses simply a balance between two extremely similar cases, and there is a certain set-off against each other of the objectionable assumptions on each side. Had one *set* of experiments only been proposed, and had we been asked to evaluate the probability of continued repetition of them confirming their verdict, I should have felt all the scruples I have already mentioned. But here we have got two sets of experiments carried on under almost exactly similar circumstances, and there is therefore less arbitrariness in assuming that their unknown conditions are tolerably equally prevalent. (*Logic of Chance*, Edn. 1888, p. 187.)

The writer seems here for a moment to realise that nearly our whole conduct in life depends on judging from samples as to sampled populations*. If past experience is no guide to future happenings, no man could get on in life, and the fact that every insurance company does judge from past experience to future happenings, and succeeds in so doing, seems to preclude us from describing all hypotheses by which we pass from one to the other as "obtaining solutions about matters on which the human intellect has no right to any definite quantitative opinion." It is difficult to understand why if we may not apply it to one population we are justified or partially justified in applying it to *two* in the above case of medical treatment. Dr Venn seems to suggest that any hypothesis applied to *both* would lead to the same result. I see no reason for this whatever. I think the true answer is this: that we have learnt by wide experience that a sample does within broad limits describe the population sampled. Otherwise the sale of corn and coal and oil in bulk would be impossible. We know further from experience that the larger the sample the more accurately it describes the population sampled. These are not hypotheses, but the conclusions of wide human experience, without which not only the work of every practical man but of every scientific investigator would be idle. The question then is whether we shall assume the sample actually describes the population sampled†—although we know by experience that it actually does not—or consider all the possible populations from which the sample might have been drawn. If we adopt the latter course, we can certainly ask whether some of these populations are more likely to have occurred *on the present occasion* than others, and weight those populations accordingly. But to do this is not to appeal to some extraneous knowledge, any such ought *à priori* to be introduced into the statement of the problem. Such extraneous knowledge usually amounts to stating that we have directly or indirectly sampled this or a like population previously, and accordingly is really equivalent to an increase in the size of the sample and can be allowed for in that way. Given the possible populations and no extraneous knowledge we have to consider how they should be most reasonably weighted to give the observed result. It is difficult to conceive any more reasonable weighting, *based on a single experimental result*, than that which weights them with the probability that if they had existed they would have

* Dr Venn would justify the comparison of two medical treatments, but would not allow an argument as to a new treatment, where there had previously existed no treatment at all, e.g. say for Sleeping Sickness, twenty years ago.

† Insurance companies usually adopt this method, and I presume justify it on the ground of the size of their samples.

given the observed result*. Had we in any particular case *knowledge* that the population sampled belonged to a particular class of populations which had marked individuals with a certain frequency distribution that would be another matter, but this knowledge is supposed excluded by the statement of the problem; we have to weight by our single sample. As I have said, such extraneous knowledge usually amounts to increasing the size of our sample. For example, if previous experience gives r' marked individuals, and s' unmarked, then $r+r'$ and $s+s'$ should be used instead of r and s, and if r'/s' is much greater than r/s there will clearly be, on the hypothesis of inverse probabilities, a great shifting of the distribution of possible populations. I have endeavoured to allow for *some à priori* knowledge of the type of populations sampled in the case of determining the most suitable value of a correlation coefficient in a sampled population, when we have some knowledge of the coefficient in similar populations (*Biometrika*, Vol. XI. pp. 352—360). But where our knowledge is based upon a single sample only, and this however large, we must either decline like Dr Venn to make any use of this knowledge—in which case practical life becomes impossible—or we must appeal to the principle of inverse probabilities to give us some measure of the accuracy with which we are describing the sampled population.

Dr Venn and some other writers assume that the possible populations might be distributed according to some other law of frequency than that demanded by the law of inverse probabilities. For example, let f_t be the frequency of populations with t marked individuals, and let $S(f_t)=1$. Then these authors next, without noticing their assumption, suppose that Nature in the particular case of the sampled population presented to us has drawn at random from this bag of sampled populations. But why should Nature have drawn *indifferently* from such a bag? She may have drawn with a bias represented by the probability q_t of coming from f_t. If so the probability of drawing a sampled population with t marked individuals is not f_t, but $f_t \times q_t$, and the only reasonable measure I personally can give to this product of two unknowns is to be based on the fact of the observed single sample and on the probability $p_{t,r}$ of drawing r marked individuals in that sample from a population with t. We thus reach $p_{t,r}=f_t \times q_t$ and we see that the weighting q_t will be directly proportional to $p_{t,r}$ and *inversely* as f_t. The dogma that the sampled population has been selected in the course of its formation from a frequency distribution $S(f_t)$ seems to me far more unreasonable than the hypothesis that we must weight our possible populations with the probabilities that they would give the observed result, the only result we know. We can if we choose appeal to the Edgeworthian doctrine† that it is a broad result of experience that populations with any number of marked individuals

* It has been objected to the use of inverse probabilities that it supposes the existence of populations which we may be fairly confident do not exist; the frequency of such populations as given by the theory will, however, be indefinitely small, and need trouble us no more than the indefinitely small frequencies of men 20 feet high or of zero stature do, when we represent by a normal curve the stature distribution of a population; or again than the infinite values of the mean and standard deviation which theoretically arise when we draw samples from a normal population.

† *Mind*, Vol. IX. pp. 223—235, 1884.

from none up to all are equally likely to be found. It is possibly a reasonable doctrine for a considerable portion of our frequency curve of populations, but I am inclined to think that the populations with none and all, giving probabilities equal to certainty, would be found to be lumped up. Let us suppose for illustration that a mark was made on our scale of skin pigmentation and the number of individuals in a particular race with a darker pigmentation was found to be r, and without knowledge of other races, we wanted to predict something about the percentage of dark individuals in the sampled population. Now clearly with such a method of estimating dark individuals in a race, the frequency distribution of populations would not be a horizontal line; there would be an excessive lumping up of the white and dark populations at either end of our curve of population frequency. It is true that some of this is effectively cut off by the knowledge that there are at least r marked and s unmarked individuals in the sampled population. But I think it more than reasonable to argue that the formation of the observed population was *not* equivalent to a random selection from the remainder $\overset{t=n-s}{\underset{t=r}{S}}(f_t)$ of such a frequency; we are justified in supposing a bias or weighting to have occurred, and the most reasonable weighting to adopt is that of choosing the frequencies to be such that they are proportional to the probability they give of the observed result. Thus I do not think it is necessary to appeal to an *à priori* equal distribution of populations, but to a reasonable weighting of any distribution, with a not necessarily random selection. The "random selection" is as much a dogma as to the formation of the sampled population as is the assumption that possible populations are *à priori* equally probable. We choose the frequency from which we are to make a random selection, not to be a straight line, but to be the one most probable on the basis of the sole event observed.

APPENDIX II.

NOTE ON A SOLUTION DUE TO LAPLACE.

Towards the end of the 18th century the question of a census of the population was becoming a prominent question in the greater European countries and an attempt to solve it for France by records of births and by test samples was made between 1783 and 1789 by illustrious scientists. In the *Mémoires de l'Académie des Sciences*, published in Paris, 1786—1791*, will be found for the following years, 1783, pp. 703—718; 1784, pp. 577—593; 1785, pp. 661—688; 1786, pp. 703—717; 1787, pp. 601—610; 1788, pp. 755—767, portions of a memoir entitled: *Essai: Pour Connoître la Population du Royaume & le nombre des habitans de la Campagne, en adaptant sur chacune des Cartes de M. Cassini, l'année commune des Naissances, tant des Villes que des Bourgs & des Villages dont il est*

* The *Mémoires* came to an end with the volume for 1789, issued in 1791, and were not resumed until after the Reign of Terror, Robespierre having proclaimed that there was no further need for men of science.

fait mention sur chaque Carte; présenté à l'Académie. Par M^rs du Séjour, le Marquis de Condorcet & de La Place.

The object of the *Essai* was to provide an estimate of the population of France and its geographical distribution. The method employed does not get except, perhaps, in accuracy of estimate, beyond that of old John Graunt; namely, it consists in determining for a sample the ratio of births in a year to the population producing those births, and then ascertaining the number of births in a year in each urban and rural district in France and multiplying by the ratio of population to births determined for the sample. This ratio was taken as 26 for the towns and villages but as 30 for Paris. It will be clear that we have here a definite case of proceeding from a sample to the population sampled. There would be little interest now in this attempt to take a census of France on John Graunt's lines were it not that a memoir by M. de La Place is affixed to the first issue of the joint memoir. It occupies pp. 693—702 of the *Mémoires* for 1783, and is entitled: *Sur les Naissances, les Mariages et les Morts à Paris depuis* 1771 *jusqu'en* 1784*; & dans toute l'étendue de la France, pendant les années* 1781 *&* 1782.

The title gives no idea of the contents of the memoir, which may be defined as follows:

Supposing a statistical ratio to be determined from two quantities x_s and y_s by means of a sample from the whole, and we know one of these quantities X_w for the whole, what error are we likely to make in estimating Y_w for the whole as equal to $X_w \times \frac{y_s}{x_s}$?

Now it is clear that the variability of Y_w will depend on three things, namely, on the variability of y_s, the variability of x_s, and thirdly on the correlation of x_s and y_s in samples of the same size. That is to say, the variability of sample populations, of sample births and the correlation of the size of population with number of births.

Now how does Laplace treat this problem and obtain a measure of the error introduced by his sample into the estimated population of France?* He applies an extension of Bayes' Theorem to predict from a first sample p, q observed, what a second sample with known q' but unknown p' might produce. There seems to me here something questionable. His first sample is a sample of the *whole* population of France, it is a part of a limited population; what right has he to consider both to be samples from a third, an infinite population, and treat them as if they were independent samples? His sample of the population is roughly one-thirtieth of the total population and this in itself limits the range of values that it is possible to give to p' in the total population.

To understand this criticism it is desirable to state the practical problem with which Laplace wishes to deal. Using a factor of 26 or 26·5, as the case may be,

* He uses the same method with a slightly different notation in his *Théorie Analytique des Probabilités*, pp. 391—394. Ed. Paris, 1812.

ascertained from the sample population as the ratio of population to annual births, we determine the total population of France, from a knowledge of the 973,054 annual births, to be 25,299,417 or 25,785,941 according to the factor used. Laplace asks what must be the size of the sample in order that there may be a very great probability that we are not more than 500,000 persons wrong in our estimate of the total population. Laplace—like many of his countrymen—seeks for a corresponding urn problem as a means of finding the required solution. He writes (p. 696) in effect as follows:

Let us consider an urn which contains an infinite number of white and black balls, and suppose that in a first drawing we have obtained p white and q black balls. Next let us suppose that in a second drawing we have obtained q' black balls, but that we do not know the number p' of white balls drawn with them; the method which presents itself for determining this number in an approximate manner is to suppose it in the same relation to q' as p to q, which gives us: $p' = pq'/q$ for the number of white balls. Accordingly we shall determine the probability that the true unknown number lies between the limits $p\frac{q'}{q}(1-w)$ and $p\frac{q'}{q}(1+w)$, or what is the same thing that the error of the result pq'/q does not exceed $p\frac{q'}{q}w$.

But it appears to me that Laplace's urn statement does not in the least fit his actual problem. We are not taking a second sample from an infinite population. We have only one sample and we want to learn something about the population from which it has been sampled, which is finite in extent, although its extent is unknown. We do know, however, that it contains q' white balls, i.e. births in all France. There appears, however, to be a slight difference between the treatment in the *Mémoires* and in the *Théorie des Probabilités*. The latter appears to deduct the births from the population, and treat the remainder as the varying quantity; in the former it appears to be the total population without deduction of births*. Whichever may be supposed taken, let us term them p_1 and p_1'.

Now Laplace takes an urn in which there are an infinite number of black and white balls and puts the chance of drawing a white ball as x. This he is going to consider as the chance of a birth, and the chance of q' births will be $x^{q'}$. Now if for a moment we forget his urn and suppose the chance of an additional unit of population to be y, the chance of drawing p_1' units will be $y^{p_1'}$, and if we suppose the units of birth and population to come in any order then the chance of q' births and p_1' population will be

$$\frac{\lfloor q' + p_1'}{\lfloor q' \, \lfloor p_1'} x^{q'} y^{p_1'}.$$

* Cf. pp. 697—698 of the *Mémoires* with p. 392 of the *Théorie*, but in both cases I venture to think the treatment obscure.

Laplace, however, following up his own problem, puts $y = 1 - x$ and says that

$$\frac{\lfloor q' + p_1'}{\lfloor q' \, \lfloor p_1'} x^{q'} (1-x)^{p_1'}$$

is the chance that, if the chance of a birth in a year be x, the ratio of births to population will be q'/p_1'. I am quite unable to follow this argument; it appears to me that the great master by adopting an urn problem has lost touch with reality. I fail to see how births and numbers in a population correspond to black and white balls in definite proportions in an urn, so that if on "trial" (trial of what?) a birth does not occur, then another unit of population must occur. It sounds almost a self-contradiction, whether we are dealing with a population which includes or excludes births. An increase of births seems to be correlated with an increase of population and not a decrease, as Laplace's urn problem suggests. Yet there is a very real problem behind Laplace's analysis and it is easy to see cases to which it might apply, but an essential condition of them is that the two variates should be *exclusive**, and this appears to be far from the truth in births and units of population.

Laplace, having reached the expression

$$\frac{\lfloor q' + p_1'}{\lfloor q' \, \lfloor p_1'} x^{q'} (1-x)^{p_1'}$$

for the chance of births q' and population p_1', now puts $p_1' = 0, 1, 2 \ldots \infty$, and sums for these values, i.e. he assumes all values of p_1' from 0 to ∞ equally likely initially, or "he distributes equally his ignorance." It is easy to see that the sum

$$= x^{q'} (1 - (1-x))^{-(q'+1)} = \frac{1}{x}.$$

Hence, applying inverse probabilities,

$$\frac{\lfloor q' + p_1'}{\lfloor q' \, \lfloor p_1'} x^{q'+1} (1-x)^{p_1'}$$

is the chance of q' being associated with p_1'. Again I think Laplace is at fault, I do not see how it is possible to give p_1' this system of values in his actual problem, because we know from our sample that there is at least a population of p, for we have sampled that number, or if we subtract births $p - q$. On Laplace's assumption that we are taking a second sample from an infinite population, this might be true. Again it is difficult to follow how a zero population could be associated with q' births!

Returning to Laplace, he now says that the probability of the ratio x occurring as judged by the first sample is (by the law of inverse probabilities)

$$\frac{\lfloor p_1 + q}{\lfloor p_1 \, \lfloor q} x^q (1-x)^{p_1} \Big/ \frac{\lfloor p_1 + q}{\lfloor p_1 \, \lfloor q} \int_0^1 x^q (1-x)^{p_1} dx = x^q (1-x)^{p_1} \Big/ \int_0^1 x^q (1-x)^{p_1} dx.$$

* For example, if every person vaccinated were registered, then Laplace's problem would apply to determining the total population from the total registered vaccinated on the basis of counting vaccinated and unvaccinated in a sample.

Thus he reaches the following expression for the combined probability of q' and p_1' occurring with the value x:

$$\frac{\lfloor p_1' + q'}{\lfloor p_1' \lfloor q'} \frac{x^{q+q'+1}(1-x)^{p_1+p_1'}}{\int_0^1 x^q (1-x)^{p_1} dx}.$$

The next step taken is to sum this for all values of p_1' from 0 to some integer s, and then to integrate for x from 0 to 1, and we thus reach the following expression for the chance of p_1' lying between 0 and s:

$$P = \frac{\int_0^1 x^{q+q'+1}(1-x)^{p_1}\left(1+(q'+1)(1-x)+\ldots+\frac{(q'+1)\ldots(q'+s)}{s!}(1-x)^s\right)dx}{\int_0^1 x^q(1-x)^{p_1}dx} \quad \ldots\ldots\ldots\ldots\text{(xxiv)}.$$

Laplace now says that when q' and s are *large*

$$1+(q'+1)(1-x)+\frac{(q'+1)(q'+2)}{2!}(1-x)^2+\ldots+\frac{(q'+1)\ldots(q'+s)}{s!}(1-x)^s = \frac{1}{x^{q'+1}}\frac{\int_{1-x}^1 x'^s(1-x')^{q'}dx'}{\int_0^1 x'^s(1-x')^{q'}dx'}.$$

This can easily be shown to be the case whatever q' and s may be, whether large or not. Hence

$$P = \frac{\int_0^1 x^q(1-x)^{p_1}\int_{1-x}^1 x'^s(1-x')^{q'}dx'\,dx}{\int_0^1 x^q(1-x)^{p_1}dx \times \int_0^1 x'^s(1-x')^{q'}dx'} \quad \ldots\ldots\ldots\ldots\text{(xxv)}.$$

Laplace then appeals to a previous *Mémoire* in which he has treated the evaluation of the integral in the numerator when q, p_1, s and q' are large numbers*. This really amounts to replacing the integral of the series for P by the integral of the best fitting normal curve. I think we can get a more exact result, and deduce Laplace's value from it, by regarding the numerator of P in (xxiv) as a series of complete B-functions:

$$P = S_{t=0}^{t=s}\frac{\lfloor q'+t}{\lfloor q' \lfloor t}\frac{B(q+q'+2,\, p_1+t+1)}{B(q+1,\, p_1+1)}$$

$$= \frac{\lfloor q+q'+1}{\lfloor q' \lfloor q+q'+p_1+2}\frac{\lfloor q+p_1+1}{}F_s(q'+1,\, p_1+1,\, q+q'+p_1+3),$$

where F_s denotes the sum of the first $s+1$ terms of a hypergeometrical series. This series is convergent because $\gamma-\alpha-\beta-1=q$. Hence by using the formulae of the paper to which this is appended we can get the sum of any number of terms by fitting with the appropriate skew curve and finding its area.

* "Sur les approximations des Formules qui sont fonctions de très-grands Nombres." *Mémoires de l'Académie des Sciences*, 1783, p. 430. Paris, 1786.

We have

$$\text{Mean} = \frac{\alpha\beta}{\gamma-\alpha-\beta-1} = \frac{(q'+1)(p_1+1)}{q} = \frac{q'p_1}{q},$$

if as Laplace supposes we are dealing with large numbers. Again

$$\sigma^2 = \frac{\alpha\beta(\lambda+\alpha)(\lambda+\beta)}{\lambda^2(\lambda-1)} = \frac{(q'+1)(p_1+1)(q+q'+1)(q+p_1+1)}{q^2(q-1)}$$

$$= \frac{q'p_1(q+q')(q+p_1)}{q^3},$$

if the numbers be large. Again,

$$\text{Mode} = \frac{q'p_1}{q+2} = \frac{q'p_1}{q} = \text{Mean},$$

if the numbers be large. Further, $\beta_1 \to 0$ and $\beta_2 \to 3$, if the same condition holds.

Accordingly the hypergeometrical series approaches the normal distribution, and this is what Laplace's method really amounts to. He takes the distribution of p_1' to be a normal curve about the value p_1q'/q as mean with the standard deviation

$$\sigma = \sqrt{\frac{q'p_1(q+q')(q+p_1)}{q^3}} \quad \text{.....(xxvi).}$$

This is in agreement with Laplace's solution in the *Mémoires de l'Académie*, if we put p_1 = the population, and in accordance with the *Théorie des Probabilités* if we put $p_1 = p - q$, i.e. population minus births in the sample. See p. 393 of the 1812 edition. Thus the problem is reduced to a determination of what the size of p_1 should be in order that

$$\left(p_1' - p_1\frac{q'}{q}\right)\Big/\sigma$$

should provide a probability, as given by the Tables of the Probability Integral, which will be less than a fixed amount. Laplace takes the case of an error of half a million in the estimated population either in excess or defect to have odds of 1000 to 1 against it. Accordingly, $500{,}000/\sigma$ must give a tail area of ·000499, or $\frac{1}{2}(1+\alpha)$ must in the usual notation be about ·9995. This connotes $\sigma = 500{,}000/3{\cdot}29$ nearly, and with the above value of σ

$$\left(\frac{500{,}000}{3{\cdot}29}\right)^2 = q'\frac{p_1}{q}\left(1+\frac{p_1}{q}\right)\left(1+\frac{q'}{q}\right).$$

Laplace puts $q' = 973{,}054{\cdot}5$ as the annual number of births in France*. He assumes on the basis of previous samplings that p_1/q is in the neighbourhood of 26; he actually takes 25·5, 26 and 26·5. If we adopt 26, we have

$$\left(\frac{500{,}000}{3{\cdot}29}\right)^2 = 973{,}054 \times 26 \times 27\left(1+\frac{q'}{q}\right),$$

whence I find $$q' = 32{\cdot}8124q,$$

and $$p_1 = 26q = 26 \times 973{,}054/32{\cdot}8124 = 771{,}032.$$

* Note the ·5, very characteristic of Laplace, especially when his whole process of representing the hypergeometrical series is only a relatively rough approximation.

Laplace gives $p_1 = 771{,}469$, sensibly in agreement. Thus according to him the relation of births to population ought to be determined on a sample of 700,000 to 800,000, if the odds are to be 1000 to 1 that we have not an error of 500,000 in the total estimated French population.

Adopting the reasoning of the *Mémoire*, in which p_1 and p_1' represent populations and q and q' births, I can see no justification for Laplace's method of reducing the problem to an urn problem. I see no reason why an additional birth in the sample means one fewer member of the population. I see further no ground whatever for considering the first sample and France as a whole as independent samples from an indefinitely large population. If we take $p_1 = p - q$ as is done in the *Théorie*, then I think the reasoning must be of this kind: the population consists of "new-borns" (as indicated by the births) and "old-borns," and when we pick a sample of a definite size a new-born may be considered to exclude an old-born and *vice versa*. But although there is a reason here—and there seemed none in the previous case—I am doubtful of its validity. It assumes that all correlation between new-born and old-born is due to sampling errors, and that there is no organic relation between the two; that is to say that we suppose in any sample that the number of new-borns is independent of the constitution of the population of old-borns, for example of their age distribution, which would vary from sample to sample, even if the samples were truly random.

Whether Laplace's Problem can or cannot be applied to the new-born considered as marked individuals in a population, it appears desirable to obtain a solution of it and compare it with Laplace's. I state it as follows:

Laplace's Problem. A population of unknown size N is known to contain q' affected or marked members. It is desired to ascertain—on the hypothesis of inverse probabilities—a measure of the error introduced by estimating N to be $n \times \frac{q'}{q}$, where q is the number of marked individuals in a sample of size n.

It is clear that the problem is a very important one and it occurs repeatedly when arguments from samples are applied to populations. If p' and p be the number of unaffected in the population N and its sample n, then by Equation (iii) above the chance $C_{p,p'}$ of a sample from N containing p unaffected individuals is

$$C_{p,p'} = \frac{\lfloor n}{\lfloor q \rfloor p} \frac{\lfloor p'}{\lfloor p' - p} \frac{\lfloor q'}{\lfloor q' - q} \frac{\lfloor N - n}{\lfloor N} \quad \text{.....(xxvii).}$$

Here n, p, q and q' are really known and $p' = N - q'$, hence

$$C_{p,p'} = \frac{\lfloor n}{\lfloor q \rfloor p} \frac{\lfloor q'}{\lfloor q' - q} \times \frac{\lfloor N - q' \rfloor N - n}{\lfloor N - q' - p \rfloor N}.$$

Now by the rule of inverse probabilities, the chance of the sampled population being N is proportional to $C_{p,p'} = C_{p,p'}/S_{p'}(C_{pp'})$. But since a sample n has been tested and found to contain p unmarked individuals and N is known to contain q'

marked individuals, N can only vary from $p+q'$ to ∞. Hence the frequencies of the various possible values of N are proportional to the terms of the series

$$1+\frac{(p+1)(q'-q+1)}{1!\,(p+q'+1)}+\frac{(p+1)(p+2)(q'-q+1)(q'-q+2)}{2!\,(p+q'+1)(p+q'+2)}+\text{etc. ....}$$

This is a hypergeometrical series of which the sum is

$$\frac{\lfloor q-2 \rfloor \lfloor p+q'}{\lfloor q'-1 \rfloor \lfloor n-1}.$$

Thus the distribution of N per mille of possible values of N is

$$\frac{1000\,\lfloor q'-1 \rfloor \lfloor n-1}{\lfloor q-2 \rfloor \lfloor p+q'}\left(1+\frac{(p+1)(q'-q+1)}{1!\,(p+q'+1)}+\frac{(p+1)(p+2)(q'-q+1)(q'-q+2)}{2!\,(p+q'+1)(p+q'+2)}\right.$$
$$\left.+\ldots+\ldots\right) \quad \ldots\ldots\ldots\text{(xxviii).}$$

Here the first term corresponds to $N=p+q'$.

This hypergeometrical can as before be very closely replaced by a Pearson Curve of which the β_1 and β_2 are given in (xviii) and (xxi) above. If we content ourselves like Laplace by replacing the hypergeometrical series by a normal curve we have

$$\text{Mean of } N=\bar{N}=p+q'+\frac{(p+1)(q'-q+1)}{q-2},$$

$$\text{Mode of } N=\tilde{N}=p+q'+\frac{p(q'-q)}{q}=n\frac{q'}{q},$$

$$\sigma_N^2=\frac{(p+1)(q'-q+1)(n-1)(q'-1)}{(q-2)^2(q-3)} \quad \ldots\ldots\ldots\ldots\text{(xxix).}$$

When p, q and q' are all large numbers, Laplace's case, then

$$N=n\frac{q'}{q}=\tilde{N},$$

or mean and mode coincide, while

$$\sigma_N^2=\frac{q'p(q'-q)(p+q)}{q^3} \quad \ldots\ldots\ldots\ldots\text{(xxx).}$$

This is clearly not in agreement with Laplace's result above (xxvi), although the *Théorie* which makes p_1 the difference of the size of the population and the number of births for the sample is more correct than the *Mémoire**. The chief

* Laplace's p should in our notation I think be $p+q$, although this is then not wholly in accordance with his mean, which he takes as $p\frac{q'}{q}$=our $p\frac{q'}{q}+q'$, $=p'+q'$, so that his p' corresponds with the whole population not new-born. I say not wholly in accord, because when he comes to actual numbers he puts $p=2{,}037{,}615$, his enumerated sample population, and not this number *less* the births. With this value for his p, Laplace's standard deviation is given by (p. 393 of the *Théorie*)

$$\sigma^2=\frac{(p+q)\,q'p\,(q+q')}{q^3}.$$

On the basis of his larger sample, Laplace makes the ratio of population to births to be 28·352,845 (!), an increase on his former 26.

difference however lies in the factor $q' - q$, where Laplace's result has $q' + q$. This difference is produced by Laplace treating his sampled population as if it were a second sample independent of that already taken. This it is not, and the two values will only approach when the size of the sample is very small as compared to that of the sampled population. In most such cases σ_N will be too large to render a determination of N trustworthy.

I venture to think, therefore, that while Laplace's Problem is most important, it does not cover the case to which he applies it, and that his solution of the problem itself is not really correct.

ON THE USE AND INTERPRETATION OF CERTAIN TEST CRITERIA FOR PURPOSES OF STATISTICAL INFERENCE. PART I.

By J. NEYMAN, Ph.D. and E. S. PEARSON, D.Sc.

CONTENTS.

		PAGE
I.	Introductory	175
II.	Sampling from a Normal Population:	
	(1) Description of the Fundamental Space	178
	(2) The Use of the Contour Systems in testing Hypothesis A	182
	(3) The Criterion of Likelihood	184
	(4) The Application of the Criterion of Likelihood in testing Hypothesis A	187
	(5) Student's Test and Hypothesis B	189
	(6) Alternative Method of examining Hypothesis B	191
	(7) Solutions obtained by the Inverse Method	192
	(8) Analysis of Church's samples from a Skew Population	197
	(9) Illustrations of the use of the Tables of P_λ	202
III.	Sampling from a Rectangular Population:	
	(10) Sampling Distributions of the Frequency Constants	208
	(11) Solutions obtained by the Inverse Method	213
	(12) Illustrations from Experimental Sampling	214
IV.	Sampling from an Exponential Population:	
	(13) Sampling Distributions of Frequency Constants	221
	(14) Test for Random Intervals	226
V.	Conclusion	230
VI.	Appendix with P_λ Tables	233

I. Introductory.

One of the most common as well as most important problems which arise in the interpretation of statistical results, is that of deciding whether or not a particular sample may be judged as likely to have been randomly drawn from a certain population, whose form may be either completely or only partially specified. We may term Hypothesis A the hypothesis that the population from which the sample Σ has been randomly drawn is that specified, namely Π. In general the method of procedure is to apply certain tests or criteria, the results of which will enable the investigator to decide with a greater or less degree of confidence whether to accept or reject Hypothesis A, or, as is often the case, will show him that further data are required before a decision can be reached. At first sight the problem may be thought to be a simple one, but upon fuller examination one

is forced to the conclusion that in many cases there is probably no single "best" method of solution. The sum total of the reasons which will weigh with the investigator in accepting or rejecting the hypothesis can very rarely be expressed in numerical terms. All that is possible for him is to balance the results of a mathematical summary, formed upon certain assumptions, against other less precise impressions based upon *à priori* or *à posteriori* considerations. The tests themselves give no final verdict, but as tools help the worker who is using them to form his final decision; one man may prefer to use one method, a second another, and yet in the long run there may be little to choose between the value of their conclusions. What is of chief importance in order that a sound judgment may be formed is that the method adopted, its scope and its limitations, should be clearly understood, and it is because we believe this often not to be the case that it has seemed worth while to us to discuss the principles involved in some detail and to illustrate their application to certain important sampling tests.

There are two distinct methods of approach, one to start from the population Π, and to ask what is the probability that a sample such as Σ should have been drawn from it, and the other the inverse method of starting from Σ and seeking the probability that Π is the population sampled. The first is the more customary method of approach, partly because it seems natural to take Π as the point of departure since in practice there are often strong *à priori* grounds for believing that this is the population sampled, and partly because there is a common tendency to view with suspicion any method involving the use of inverse probability. But in fact, however strong may be the *à priori* evidence in favour of Π, there would be no problem at all to answer if we were not prepared to consider the possibility of alternative hypotheses as to the population sampled; and we shall find that it is impossible to follow the first method very far without introducing certain ideas of inverse probability—that is to say, arguing from the sample to the population. If on the other hand we start boldly with assumptions regarding *à priori* and *à posteriori* probability, we reach by an almost simpler method sampling tests very nearly equivalent to those obtained from the first starting-point. Indeed the inverse method may be considered by some the more logical of the two; we shall consider first however the other solution.

Perhaps the most suggestive method of description is to represent Σ by a point in a hyperspace whose dimensions will depend upon the particular problem considered; and to associate the criteria for acceptance or rejection with a system of contours in this space, so chosen that in moving out from contour to contour Hypothesis A becomes less and less probable*. The frequency with which the sample corresponding to a particular point will occur in random sampling from Π may be represented by giving to the space an appropriate "point-density." Thus the chance of drawing a sample whose representative point lies within a certain

* Here and later the term "probability" used in connection with Hypothesis A must be taken in a very wide sense. It cannot necessarily be described by a single numerical measure of inverse probability; as the hypothesis becomes "less probable," our confidence in it decreases, and the reason for this lies in the meaning of the particular contour system that has been chosen.

region in the space is proportional to the density integrated throughout that region, or in other words to the "weight" of the region. The contours are not necessarily contours of equal density, but may be surfaces or regions throughout which some statistical measure such as the mean or standard deviation remains at a constant "level." Although it is impossible to visualise a space of high dimensions, it will be found that with the help of the analogy of a three-dimensional density space this method of description is often a very considerable aid to purely algebraic discussion.

Setting aside the possibility that the sampling has not been random or that the population has changed during its course*, Σ must either have been drawn randomly from Π or from Π', where the latter is some other population which may have any one of an infinite variety of forms differing only slightly or very greatly from Π. The nature of the problem is such that it is impossible to find criteria which will distinguish exactly between these alternatives, and whatever method we adopt two sources of error must arise:

(1) Sometimes, when Hypothesis A is rejected, Σ will in fact have been drawn from Π.

(2) More often, in accepting Hypothesis A, Σ will really have been drawn from Π'.

In the long run of statistical experience the frequency of the first source of error (or in a single instance its probability) can be controlled by choosing as a discriminating contour, one outside which the frequency of occurrence of samples from Π is very small—say, 5 in 100 or 5 in 1000. In the density space such a contour will include almost the whole weight of the field. Clearly there will be an infinite variety of systems from which it is possible to choose a contour satisfying such a condition. For example there will be the system of contours upon any one of which (*a*) the mean, or (*b*) the standard deviation, or (*c*) the ratio of mean to standard deviation, of Σ is constant.

The second source of error is more difficult to control, but if wrong judgments cannot be avoided, their seriousness will at any rate be diminished if on the whole Hypothesis A is wrongly accepted only in cases where the true sampled population, Π', differs but slightly from Π. It is not of course possible to determine Π', but making use of some clearly defined conception of probability we may determine a "probable" or "likely" form of it, and hence fix the contours so that in moving "inwards" across them the difference between Π and the population from which it is "most likely" that Σ has been sampled should become less and less. This choice also implies that on moving "outwards" across the contours, other hypotheses as to the population sampled become more and more likely than Hypothesis A.

* A non-random or biased sample from Π may appear to be a random sample from Π'; in practice it may be very important to be able to attribute Σ to Π rather than Π'. But clearly no general rules to cover such cases can be given, for the position must depend in each case on the possible forms of bias that may have occurred.

Both these aspects of the problem must be taken into account. Using only the first control, any given sample could always be found to lie outside an extremely divergent contour of *some* system. That is to say a criterion of any desired degree of stringency could always be found for a sample. But regarding the position from the second point of view, it is seen that there will only be certain systems which are of any value as criteria.

The application of these principles will become clearer when illustrated in the cases of particular sampling tests, but it seems well to emphasise at the outset the importance of careful thinking in these matters. For example it might readily be supposed that if, in sampling from Π, a sample of form Σ_1 occurs more frequently than one of form Σ_2, then greater confidence in Hypothesis A would be justified if Σ_1 were drawn rather than Σ_2. But we shall see that under certain conditions this may not be the case, because in the first event alternative hypotheses are relatively far more probable than A than they are in the second. It is indeed obvious, upon a little consideration, that the mere fact that a particular sample may be expected to occur very rarely in sampling from Π would not in itself justify the rejection of the hypothesis that it had been so drawn, if there were no other more probable hypotheses conceivable.

II. Sampling from a Normal Population.

(1) *Description of the Fundamental Space.*

We shall discuss first the case of sampling from an "infinite" population in which the single variable under consideration follows a Normal Distribution, and we shall suppose that this population, Π, is completely specified, its mean a and standard deviation σ being known. We have a sample, Σ, of n observations, $X_1, X_2, \dots X_n$, and the hypothesis whose probability we wish to test is that Σ is a random sample from Π; this is Hypothesis A. The methods to be discussed are perfectly general, but we have particularly in mind the case of small samples where the data are insufficient for the application of the ordinary (P, χ^2) test for goodness of fit.

Making use of the geometrical method first introduced into this problem by R. A. Fisher*, we shall imagine an n-dimensioned space in which we take an origin at the point $O\ (a, a, \dots a)$ and rectangular axes $Ox_1, Ox_2, \dots Ox_n$. Referred to these axes, Σ is the point $x_1 = X_1 - a, x_2 = X_2 - a, \dots x_n = X_n - a$. We now fill this space with the density field appropriate for samples of n drawn from Π, making the point-density at $(x_1, x_2, \dots x_n)$ equal

$$D = 1/(\sqrt{2\pi}\sigma)^n \,.\, e^{-\frac{1}{2\sigma^2}(x_1^2 + x_2^2 + \dots + x_n^2)} \qquad \dots\dots\dots\dots\dots\dots\text{(i)},$$

so that the chance of drawing a sample with variates lying within the range $x_1 \pm \frac{1}{2}dx_1, \dots x_n \pm \frac{1}{2}dx_n$ is $Ddx_1, \dots dx_n$.

* *Biometrika*, Vol. x. p. 507.

In the great majority of problems we are neither able nor do we wish to distinguish between $x_1, x_2, \ldots x_n$ as differentiated variables, but in the space as defined there will be $n!$ points corresponding to a given unordered set of n values of the variable. Each of these points will lie in a different region of the general hyperspace; one such region will be that for which

$$x_1 \leqslant x_2 \leqslant \ldots \leqslant x_n \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(ii)},$$

which may be described as a wedge-shaped region lying between $n-1$ primes passing through the line

$$x_1 = x_2 = \ldots = x_n \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(iii)}.$$

These $n!$ regions are identical in shape and density distribution and one may be superposed upon another or all combined together by a rotation about the axis (iii)*. The contour surfaces of equal density and the other contours that will be considered are all figures of revolution about this axis, and it will follow that any conclusions which may be drawn regarding the integral of the density taken throughout regions lying between such contours will apply to a single

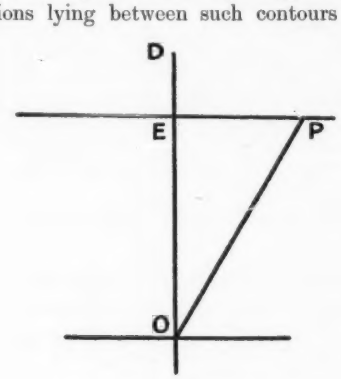


FIG. 1

segmental region as well as to the complete space. For simplicity in treatment we shall therefore consider the latter space, which may be termed the *fundamental space*, bearing in mind that in practice it is generally only possible to locate Σ in a single region such as that defined by the conditions (ii). This region may be supposed filled with a density obtained by superposing the $n!$ similar regions.

Denote the mean of Σ by $a+m$ and its standard deviation by s. The section of the fundamental space by the two-dimensioned plane which passes through the line (iii) and the point P representing Σ is shown in Figure 1. O is the origin, OD the line (iii), and E the point where this line is cut by

$$x_1 + x_2 + \ldots + x_n = nm \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(iv)}.$$

* In the case $n=2$, the regions are the halves of the x_1, x_2 plane divided by the line $x_1=x_2$. For $n=3$, there are six regions, each being a "wedge" or "slice" in three-dimensioned space lying between two planes which cut at an angle of 60° in the line $x_1=x_2=x_3$. We shall use the term "prime" to describe any linear function of the n variables. Such a locus is a "flat" space of $n-1$ dimensions lying in n dimensions.

The prime (iv) passes through P and lies at right angles to OD, so that the angle PEO is a right angle. Further P is the point $(X_1 - a, X_2 - a, \ldots X_n - a)$ and E the point $(m, m, \ldots m)$, so that

$$OP^2 = \sum_{t=1}^{n} (X_t - a)^2 = n(m^2 + s^2),$$

$$OE^2 = nm^2 \text{ and consequently } EP^2 = ns^2.$$

Hence if we reduce the scale of this section of the hyperspace in the linear ratio of $1/\sqrt{n}$, we shall obtain a plane in which the position of the point corresponding to P is given by rectangular coordinates (m, s). This plane will be termed the (m, s)-plane. Any contour in the hyperspace, whose equation can be expressed in terms of the two variables m and s only, may therefore be obtained by a rotation of the corresponding curve in the (m, s)-plane about the axis (iii), and a linear enlargement in the ratio $\sqrt{n} : 1$.

Consider now the following:

(a) The point density D in the hyperspace may be written

$$D = \text{constant} \times e^{-\frac{n(m^2+s^2)}{2\sigma^2}} = \text{constant} \times e^{-\frac{1}{2}\chi^2},$$

where*

$$\chi^2\sigma^2 = \sum_{t=1}^{n} (X_t - a)^2 = n(m^2 + s^2).$$

(b) The contours of equal density are $(n-1)$-fold hyperspheres centred at the origin O.

(c) If the two-dimensioned (m, s)-plane be rotated in m dimensions about the axis (iii), the point P will trace out an $(\eta - 2)$-fold hypersphere lying in a prime perpendicular to the axis of rotation; the boundary surface of this hypersphere is proportional to s^{n-2}.

Consequently the integral of D taken between any contour surfaces in the fundamental space which can be obtained (after enlargement) by a rotation of curves lying in the (m, s)-plane, will be the same as the integral of d between the corresponding curves in that plane, where d is a point-density assigned to the plane measured by

$$d = \text{constant} \times s^{n-2} e^{-\frac{n(m^2+s^2)}{2\sigma^2}},$$

or choosing the constant so that

$$\int_{-\infty}^{+\infty} \int_{0}^{\infty} d \,.\, dm \,.\, ds = 1,$$

$$d = \frac{n^{\frac{n}{2}}}{\sqrt{\pi}\, 2^{\frac{n-2}{2}}\, \Gamma\left(\frac{n-1}{2}\right)} \frac{s^{n-2}}{\sigma^n} e^{-\frac{n(m^2+s^2)}{2\sigma^2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(v)}.$$

* This χ^2 must not be confused with the χ^2 of the Goodness of Fit Test.

The equations of the five contour systems defined below can all be expressed in terms of m and s. In order to grasp their bearing on the problem of testing Hypothesis A, it is necessary to consider their form and position in the fundamental space, but their reduction to the (m, s)-plane makes clearer their relation to one another and enables a ready integration of D to be obtained in certain regions of the hyperspace.

These contours, which are shown diagrammatically in the (m, s)-plane in Figure 2, are as follows:

(1) The contours of equal density D (contours of χ). These are semicircles, centred at O, of radius $\chi\sigma/\sqrt{n} = \sqrt{m^2 + s^2}$.

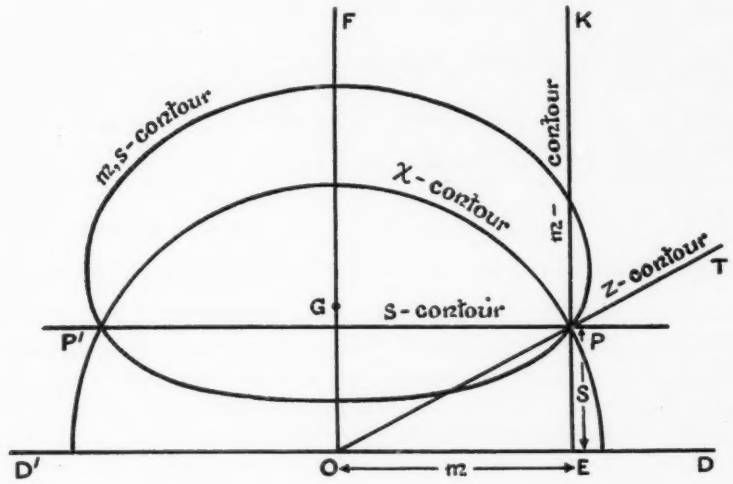


FIG. 2

(2) The contours of constant m; a series of lines parallel to the axis of s.

(3) The contours of constant s; a series of lines parallel to the axis of m.

(4) The contours upon each of which the ratio of $z = m/s$ is constant (contours of z). These are a series of straight lines radiating from O.

(5) The contours of equal density d, or of equi-probable doublets (m, s). They form a series of oval curves in the (m, s)-plane with equations

$$(n-2)\log_e s - \tfrac{1}{2}n(m^2 + s^2)/\sigma^2 = K \quad \text{.....(vi)},$$

and surround the point of maximum density G, or $m = 0$, $s = \sigma\sqrt{(n-2)/n}$.

By integrating d outside the first four contours in the (m, s)-plane we obtain the four standard integrals for the chance, P, of drawing a sample with χ, m, s, or z, equal to or greater than the value observed:

(1) $$P_\chi = c_1 \int_\chi^\infty \chi^{n-1} e^{-\frac{1}{2}\chi^2} d\chi \quad \text{.....(vii),}$$

(2) $$P_m = c_2 \int_m^\infty e^{-\frac{nm^2}{2\sigma^2}} dm \quad \text{.....(viii).}$$

(3) The distribution of s being skew, we may take

$$P_s = c_3 \int_s^\infty s^{n-2} e^{-\frac{ns^2}{2\sigma^2}} ds \quad \text{.....(ix),}$$

or

$$P_s' = c_3 \int_0^s s^{n-2} e^{-\frac{ns^2}{2\sigma^2}} ds \quad \text{.....(x),}$$

according as s is greater or less than the modal value $s_0 = \sigma\sqrt{(n-2)/n}$.

(4) $$P_z = c_4 \int_z^\infty (1+z^2)^{-\frac{n}{2}} dz \quad \text{.....(xi),}$$

and finally, for the chance of obtaining a less probable doublet (m, s),

(5) $$P_{m,s} = c_5 \int_\Omega s^{n-2} e^{-\frac{n(m^2+s^2)}{2\sigma^2}} dm\,ds \quad \text{.....(xii),}$$

where $\int_\Omega$ is to be taken outside the contour of system (vi), which passes through the sample point. This integral has not so far been computed.

The form of the contours in the fundamental space obtained by rotating Figure 2 about $D'OD$ in n dimensions can be easily visualised in the case of $n = 3$. The χ-contour becomes a sphere, the m-contour a plane at right angles to the axis, the s-contour a cylinder, the z-contour a cone, and the (m, s)-contour a hollow ring.

The foregoing results are not of course new, but this restatement has seemed necessary in order to make clear the discussion which follows*.

(2) *The Use of the Contour Systems in testing Hypothesis A.*

In the first place it is clear that any of the five contour systems described above will enable us to control the first source of error referred to in the introductory section. By rejecting Hypothesis A only when the point representing Σ falls upon a contour for which P (whether P_χ, P_m, ...) is, say, $\leqslant \cdot 01$, we shall be certain in the long run, in random sampling, of only rejecting the hypothesis when Σ has in fact been drawn from Π, in 1 case out of 100. But the samples rejected and those included will vary according to the system chosen. Consider the systems individually.

* See, for example, Student, *Biometrika*, Vol. VI. p. 1, 1908; R. A. Fisher, *Biometrika*, Vol. X. p. 507, 1914; *Metron*, Vol. V. No. 3, p. 3, 1925.

(2), (3) and (4), contours of m, s and z.

If we wish to judge whether a certain deviation in the mean, or a certain standard deviation, or a ratio of the two, is likely to be found in sampling from Π, these systems are entirely satisfactory. For instance, the smaller be P_m, the more divergent is the sample mean at $a+m$ from the population mean at a, the less frequently will a sample with mean in the limits $a+m \pm \frac{1}{2}dm$ be drawn in random sampling, and other things being equal the more likely become alternative hypotheses as to the value of the population mean.

But if our purpose is to make full use of the information supplied by Σ, and to ask whether, taken as a whole, it is likely to have been drawn from Π, the contours of m, s, and z are inadequate. In the first place they are not closed contours, for the space bounded by any one of them stretches away to infinity; this is true of the fundamental space as well as of the (m, s)-plane. Consequently the space on the inner side of any contour, however near to unity be the P associated with it, will contain points representing samples which are infinitely divergent from the population type. Therefore in using one of these systems alone as a criterion we should sometimes find ourselves being led to accept samples for which our common sense tells us that Hypothesis A is extraordinarily improbable.

It will also be seen that, while in the fundamental space the contours of the criteria of (viii) and (ix) for m and s contain the central region of the density field, the z-contours are orthogonal to the surfaces of equal density, and select out for rejection in a quite arbitrary fashion, the sample points lying at right angles to the axis (iii)*.

(1), contours of χ†.

These contours alone have the property of coinciding in the fundamental space with the surfaces of equal density, so that any contour divides off samples of more from those of less frequent occurrence. P_χ will provide a satisfactory measure in any problem where the judgment of the statistician is aided by a scale which expresses in numerical terms the fact that, in drawing from a given population, a sample is of relatively common, rather rare, exceptionally rare occurrence, etc. But in the great majority of problems we cannot so isolate the relation of Σ to Π; we reject Hypothesis A not merely because Σ is of rare occurrence, but because there are other hypotheses as to the origin of Σ which it seems more reasonable to accept. This attitude may be illustrated by the following example.

Suppose that Π is "normal" with mean at the origin of reference and $\sigma = 100{\cdot}0$, and that Σ is a sample of 8 for which the variate values are as follows:

$$+1{\cdot}0,\ +0{\cdot}4,\ -0{\cdot}6,\ +0{\cdot}1,\ -1{\cdot}1,\ -0{\cdot}3,\ -0{\cdot}1,\ +0{\cdot}2.$$

* This criticism of the z-contours does not of course apply to their use in Student's Problem, where the standard deviation of Π is unknown, which is discussed below. In the case where σ is known, however, they provide a good illustration of a system which gives for any sample a measure P_s, which is completely useless as a criterion to apply in testing Hypothesis A.

† The χ-contours here considered are of course quite different from those occurring in the (P, χ^2) test for goodness of fit, or what may be termed the "χ^2 group test." Here the variables are $x_1, x_2, \ldots x_n$; there they are the group frequencies $n_1, n_2, \ldots n_k$.

Here $$\chi^2 = \sum_{t=1}^{t=8} (X_t - a)^2/\sigma^2 = \cdot 000288.$$

P_χ is therefore very close to unity, the sample point lying almost at the centre of the density field. Yet unless the investigator had very strong *à priori* grounds for supposing that the standard deviation of the sampled population was 100·0, he would instinctively reject Hypothesis A, feeling that some other hypothesis, that gave σ in the neighbourhood of 0·5, was far more probable. This would appear to be a decision of common sense; we shall seek later to express it in more exact terms.

(5), the (m, s)-contours or contours of constant d.

These, just as the χ-contours, have the advantage of being closed, and if we take as the criterion of rejection $P_{m,s} \leqslant \cdot 01$, let us say, we shall not accept samples for which Hypothesis A seems exceedingly improbable on any common-sense grounds. As only the most divergent of the contours approach the origin O, we should reject Hypothesis A in the case of the numerical example given above. Further, since in sampling from a Normal population the mean and standard deviation may be considered as the two most important descriptive constants, the (m, s)-plane has a certain fundamental importance and it seems rational to adopt the contours of equal density in that plane. At any rate in making use of them we are applying what may be termed a "two constant" test, as distinct from the less adequate "one constant" test of either m or s alone*. But it must be remembered that it is difficult to find any logical reason for accepting the contours of equi-probable doublets (m, s), rather than those of the doublets (m, s^2), or (m, s^2), etc. For a given sample P_{m,s^2} will differ from $P_{m,s}$; the contours of the former in the fundamental space are obtained by rotating the oval curves in the (m, s)-plane for which

$$(n-3)\log_e s - \tfrac{1}{2}n\,(m^2 + s^2)/\sigma^2 = K \quad \text{.....(xiii)},$$

an equation differing from (vi) in the factor $n-3$ instead of $n-2$.

We must in fact consider whether it is possible to find a contour system which will take into account the probability of alternative hypotheses, and will not depend for its validity on the particular statistical constants chosen to describe the sample.

(3) *The Criterion of Likelihood.*

Suppose there to be two hypotheses regarding the population from which a given sample has been drawn; there will be one density field for Hypothesis A corresponding to Π, and another for Hypothesis A′ corresponding to Π'. By surrounding the sample point in the fundamental space with an element of volume $dX_1, dX_2, \ldots dX_n$ we have in the ratio of the "weights" of the elements a measure of the relative frequency of occurrence, on the two hypotheses, of samples lying

* It will be seen that if both the m-test and the s-test are applied to a sample, and the Hypothesis A is rejected if, let us say, $|m| > 3\sigma_m$ and $|s - \sigma| > 3\sigma_s$, the criterion of rejection will form a rectangle in the (m, s)-plane corresponding very roughly to the oval of the (m, s)-contours.

within the prescribed limits. These elements may seem a little artificial since we know in fact that the n variates of the sample have each unique values, even if these cannot be determined exactly. It will therefore be well to consider the position more carefully.

The possible range of the variable $x = X - a$ may be broken up into small subranges of equal length h. The n variates of the sample will each fall into one of these subranges, and the corresponding point P in the fundamental space will fall into one of the n-dimensioned hypercubes of side h which make up the framework into which the space will now be divided. As pointed out above, on p. 179, for a given set of unordered values $x_1, x_2, \ldots x_n$ there will be a large number of hypercubes in the complete space into which P may fall. If n_1 of the variates fall into one subrange, n_2 into another and so on, where $n_1 + n_2 + \ldots + n_k = n$, then the chance of drawing a sample with its representative point in one of these cubes is

$$C_p = \frac{n!}{n_1!\, n_2! \ldots n_k!}\, p \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(xiv)},$$

where p is the "weight" of one of the cubes. If none of the variates falls into the same subrange this expression reduces to $n!\, p$.

If we confine our attention to a single one of the $n!$ segmental regions such as that defined by (ii), we may consider the form of (xiv) as due to a boundary effect. When several of the x's fall into a single subrange, only a portion of the corresponding hypercube or $1/(n_1!\, n_2! \ldots n_k!)$ lies within the region (ii), part being cut off by one or more of the $n-1$ primes bounding the region.

As h/σ is decreased so as to make the block surrounding P a differential element of volume, this boundary effect becomes negligible, none of the n variates falling into the same differential subrange. On the other hand as h/σ is increased the effect becomes more and more important, until all the hypercubes within the field of significant density cut a considerable number of the bounding primes. In the first case we measure the chance of drawing a sample having the n variates in the range

$$x_1 \pm \tfrac{1}{2}h, \ldots x_n \pm \tfrac{1}{2}h,$$

by

$$f(x_1), f(x_2), \ldots f(x_n)\, h^n \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(xv)},$$

where for the Normal population

$$f(x) = 1/\sqrt{2\pi}\sigma \,.\, e^{-\frac{1}{2}\frac{x^2}{\sigma^2}};$$

in the second case, by a term of the multinomial expansion, namely

$$\frac{n!}{n_1!\, n_2! \ldots n_k!}\, p_1^{n_1} \ldots p_k^{n_k} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(xvi)},$$

where

$$p_s = \int_{x_s - \frac{1}{2}h}^{x_s + \frac{1}{2}h} f(x)\, dx.$$

Now it can be shown that when h is a finite subrange, a comparison of the probability of two hypotheses obtained from two expressions such as (xvi) leads

directly to the ordinary (P, χ^2) test for goodness of fit*. It would seem therefore that by supposing that each observation lies somewhere within a differential element of fixed length, and comparing expressions of form (xv), we employ a method which differs only in degree from that used in cases of broad grouping. The use of a grouping unit of fixed length† whatever be the hypothetical population is common to both methods; the difference consists mainly in the fact that in one case the grouping unit is so large that many variates will generally fall into each group, while in the other the great majority of groups are vacant, and those filled contain only one observation.

We shall therefore take the ratio of two expressions such as (xv), with a common h, that is to say the ratio of the point-densities at P in the two fields, as a measure for the comparison of Hypotheses A and A′. Probability is a ratio of frequencies and this relative measure cannot be termed the ratio of the probabilities of the hypotheses, unless we speak of probability *à posteriori* and postulate some *à priori* frequency distribution of sampled populations. Fisher has therefore introduced the term *likelihood*, and calls this comparative measure the ratio of the likelihoods of the two hypotheses‡. For a given sample and two given hypotheses, the ratio of the likelihoods is completely independent of the coordinate space in which the sample point is represented. For example the ratio of the densities of the two fields at P is the same in the space of $x_1, x_2, \ldots x_n$ as in that of $x_1^2, x_2^2, \ldots x_n^2$. This is a point of considerable importance.

As an illustration, return to the example of p. 183. We may suppose that there are two hypotheses as to the population: (A_1) Mean at O, $\sigma = 100{\cdot}0$, and (A_2) Mean at O, $\sigma = 0{\cdot}5$. Then the two corresponding densities at the sample point are (A_1) $6{\cdot}42 \times 10^{-20}$, ($A_2$) $5{\cdot}18 \times 10^{-4}$ and the ratio of the likelihood of Hypothesis A_1 to that of A_2 is $1{\cdot}24 \times 10^{-16}$. This ratio confirms the common-sense judgment that A_2 is far more plausible than A_1.

There is little doubt that the criterion of likelihood is one which will assist the investigator in reaching his final judgment; the greater be the likelihood of some alternative Hypothesis A′ (not ruled out by other considerations), compared with that of A, the greater will become his hesitation in accepting the latter. It is true that in practice when asking whether Σ can have come from Π, we have usually certain *à priori* grounds for believing that this may be true, or if not so, for expecting that Π' differs from Π in certain directions only. But such expectations can rarely be expressed in numerical terms. The statistician can balance the numerical verdict of likelihood against the vaguer expectations derived from *à priori* considerations, and it must be left to his judgment to decide

* This point will be discussed later in connection with the (P, χ^2) test. R. A. Fisher has dealt with it in considering the relation between Maximum Likelihood and Minimum χ^2, *Phil. Trans.* A, Vol. 222, p. 357; *Journ. of Roy. Stat. Soc.* Vol. LXXXVII. p. 446.

† It might be supposed that the element should be made to depend upon the parameters of the hypothetical population. But if, for example, h were taken to be proportional to σ, it is seen that the Normal population, for which the chance of drawing a sample with observations in the prescribed limits is a maximum, will have $\sigma = \infty$. Such an interpretation of the element would be clearly of little value.

‡ *Phil. Trans.* A, Vol. 222, p. 326.

at what point the evidence in favour of alternative hypotheses becomes so convincing that Hypothesis A must be rejected.

As for all criteria, there is a limitation to its value in the case of very small samples. If all hypotheses were equally probable *à priori*, the criterion provides a means of picking out the most likely ones *à posteriori*, whatever be the size of n. But if we have as is usual some *à priori* grounds for believing in A, we shall not feel disposed to give up this belief unless the alternative suggested by the method of likelihood is what may be termed a stable alternative. For example, in the previous illustration, while it seemed reasonable to reject Hypothesis A upon the evidence of a sample of 8, we should not do so if we had only a sample of 2 with values $+0{\cdot}4$ and $-0{\cdot}6$, because we should be quite ready to find on adding a third observation to the sample that it had a value of $+50{\cdot}0$ or $-100{\cdot}0$, or of almost any quantity whatsoever. Here the alternative of a population with $\sigma = 0{\cdot}5$ does not appear sufficiently stable to be very seriously considered.

(4) *The Application of the Criterion of Likelihood in testing Hypothesis A.*

We shall suppose that while there is uncertainty as to the values of the mean and standard deviation, a and σ of the population sampled, there are good reasons for believing that Σ has been drawn from *some* normally distributed population. This is not a necessary limitation of the general theory of likelihood, for it is possible, in theory at any rate, to compare the point-densities of any types of field whatsoever, but we consider here only the simplest case*. Taking $\bar{X}$ as the mean of Σ referred to a fixed origin and s as before, its standard deviation, the density at P of the field associated with some population Π', mean at a', and standard deviation σ', is given by

$$D = \frac{1}{(\sqrt{2\pi}\sigma')^n} e^{-\frac{n}{2\sigma'^2}\{(\bar{X}-a')^2+s^2\}} \quad \text{.....(xvii).}$$

To find the Π' for which D is a maximum we solve

$$\frac{\partial D}{\partial a'} = 0 \text{ or } \bar{X} = a',$$

$$\frac{\partial D}{\partial \sigma'} = 0 \text{ or } -\frac{n}{\sigma'^{n+1}} + \frac{1}{\sigma'^n}\frac{n\{(\bar{X}-a')^2+s^2\}}{\sigma'^3} = 0,$$

giving $$\sigma'^2 = (\bar{X}-a')^2 + s^2.$$

The population of maximum likelihood† is therefore that for which $a' = \bar{X}$ $\sigma' = s$, and the corresponding field gives as the density at the sample point P,

$$D_M = 1/(\sqrt{2\pi}s)^n \,.\, e^{-\frac{n}{2}} \quad \text{.....(xviii).}$$

From this we obtain the ratio

$$\lambda = \frac{\text{Likelihood of } \Pi}{\text{Likelihood of } \Pi' \text{ (max.)}} = \frac{D}{D_M} = \left(\frac{s}{\sigma}\right)^n e^{-\frac{n}{2}\left(\frac{m^2+s^2}{\sigma^2}-1\right)} \quad \text{.....(xix).}$$

* The effect of a moderate divergence from normality is considered experimentally below.

† It can be shown on a further differentiation that this is a maximum solution.

This ratio remains constant upon a certain contour surface in the density space of Π. That is to say, the ratio of the likelihood of Π to the maximum likelihood that can be associated with an alternative normal population is constant upon this contour, but steadily diminishes on passing outwards from one contour to another of the system. The surfaces are obtained by the rotation about the axis (iii) of the following curves in the (m, s)-plane:

$$(m^2+s^2)/\sigma^2 - \log_e(s^2/\sigma^2) = 1 - (2/n)\log_e \lambda = k \log_e 10 \quad \text{.....(xx).}$$

This equation represents a series of oval curves surrounding the point $m=0$, $s=\sigma$; the particular contour is determined by the constant k, and as n appears only on the right-hand side of the equation it will be seen that the form of the system as a whole is independent of the size of the sample. For large values of n the curves will differ only slightly from those of the contours of equi-probable doublets (m, s) given in (vi). In the fundamental space the surfaces will be closed contours which will not contain the origin. For the case $n=3$ it is possible to visualise the system of hollow rings obtained by rotating about DOD' curves much like the (m, s)-contour drawn in Figure 2.

Without claiming that this method is necessarily the "best" to adopt, we suggest that the use of this contour system in association with the density field of Π provides at any rate one clearly defined method of discriminating between samples for which Hypothesis A is more probable and those for which it is less probable. It is a method which takes into account the likelihood of alternative hypotheses, and while its justification does not rest on the choice of any particular statistical constants, the close agreement as n becomes large of its contour system with that of the equi-probable doublets (m, s) provides another form of interpretation.

To employ the method, it is necessary to know

$$P_\lambda = c_5 \int_\omega s^{n-2} e^{-\frac{n(m^2+s^2)}{2\sigma^2}} \, dm \, ds \quad \text{.....(xxi),}$$

an expression corresponding to (xii), except that here the integral is to be taken outside the contour of (xx) which passes through the point (m, s) corresponding in the (m, s)-plane to the observed sample, Σ. If then we only reject Hypothesis A when P_λ is, let us say, $\leqslant .01$, we can control the error of form (1)*, while the use of the contours of likelihood will minimise as far as possible the effect of errors of form (2). It would be possible to use the ratio λ of (xix) as a criterion, but this without a knowledge of P_λ does not enable us to estimate the extent of the form (1) error.

Tables of P_λ for $n=3$ to 50 are given in the Appendix. These are entered with k, which may either be calculated for the case of a given sample exactly from (xx), or found readily with sufficient accuracy for most practical purposes from the diagrams of Figures 13 and 14. The method by which the Tables were

* See p. 177.

computed and also the manner in which P_λ can be obtained for higher values of n are described in the Appendix. Illustrations of the use of the Tables are given in Section (9) below.

(5) *Student's Test and Hypothesis B.*

This is the problem to which Student's z-test has been applied*. The sample, Σ, of n observations $X_1, X_2, \ldots X_n$ has a mean $\bar{X}$ and a standard deviation s. There is every reason to believe that Σ has been drawn from *some* normally distributed population, and we wish to test Hypothesis B—namely, that Σ has been randomly drawn from such a population with mean a, but unspecified standard deviation.

In the fundamental space the axes and the position of the point P representing Σ are as before, but the density field appropriate to Π depends upon the unspecified σ. Of the contour systems considered above only those of z are independent of σ, and it might therefore appear at first sight that here was sufficient justification for the use of Student's integral, P_z, to test the probability of Hypothesis B. But B is really a multiple hypothesis concerning the sub-universe of normal populations, $M(\Pi)$, with means at a and with varying standard deviations. It only becomes precise upon definition of the manner in which σ is distributed within this sub-universe, that is to say, upon defining the *à priori* probability distribution of σ. Although it is true that whatever this may be, the distribution of z in sampling from $M(\Pi)$ will be the same with the probability integral given by (xi), it does not follow that the z-contours will therefore be the appropriate contours to use in testing the hypothesis as a whole. If it were possible to determine an *à priori* distribution for σ,

$$y = \phi(\sigma), \text{ where } \int_0^\infty \phi(\sigma) = 1 \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(xxii)},$$

then we might fill the fundamental space with a density field appropriate to $M(\Pi)$ given by

$$\Delta = \frac{1}{(2\pi)^{n/2}} \int_0^\infty \frac{1}{\sigma^n} \phi(\sigma)\, e^{-\frac{1}{2\sigma^2} S(X_i - a)^2}\, d\sigma \ldots\ldots\ldots\ldots\text{(xxiii)},$$

and consider what in this were the appropriate contours to use. In practice, however, it is doubtful if $\phi(\sigma)$ could ever be determined, and if some arbitrary form is to be assigned, it is simpler to approach the problem from the point of view of inverse probability, and think of the density space as associated with Σ rather than $M(\Pi)$. This will be done in a later section.

In a problem where Hypothesis B is to be tested, the investigator will usually be quite certain that σ cannot have all values between 0 and ∞. He may often be able to define broad limits outside which it is impossible that it should lie, and to suggest narrower but less clearly defined limits within which it most probably lies. But although he might say, "As far as I know, σ may lie equally well anywhere between b_1 and b_2," he would probably find it difficult on reflection to

* *Biometrika*, Vol. VI. p. 1 *et seq.*

justify this statement, rather than, "As far as I know, σ^2 may lie equally well anywhere between b_1^2 and b_2^2." In fact, his *à priori* information as to σ is generally not sufficiently precise to bear exact numerical expression. Under such conditions the criterion of likelihood will probably be of service.

Corresponding to any given Σ, we can find the population Π out of $M(\Pi)$ for which the likelihood is a maximum. This is obtained from (xvii) by equating $\partial D/\partial\sigma'$ to zero, considering $a = a'$ as fixed. We find $\sigma'^2 = m^2 + s^2$, and for this population the density at P is

$$D_0 = \frac{1}{(2\pi)^{n/2}} \frac{1}{(m^2+s^2)^{n/2}} e^{-\frac{n}{2}} \quad \text{.....(xxiv).}$$

We can now obtain the ratio of (a) maximum likelihood for a member of $M(\Pi)$, to (b) maximum likelihood for any normal population (the D_M of (xviii)), or

$$\lambda' = \frac{D_0}{D_M} = \left(\frac{s^2}{m^2+s^2}\right)^{\frac{n}{2}} = (1+z^2)^{-\frac{n}{2}} \quad \text{.....(xxv).}$$

This ratio is constant along the contours of z. It follows that if the sample point is surrounded by an element of volume, the ratio of the maximum frequency with which a sample lying within these limits will be drawn from a normal population with mean at a to the maximum frequency with which it could be drawn from any unrestricted normal population whatever, is constant upon one of these contours. Speaking therefore in somewhat loose terms, we may say that, other things being equal, we shall be more ready to accept Hypothesis B for a sample lying inside or on a certain z-contour, than for one lying outside it. And we may measure our confidence in a decision to reject or accept the hypothesis by P_z, which is the integral of the density outside the contour whatever be σ. Student's test was originally devised to allow for the use of s instead of σ in testing the significance of a deviation in the sample mean. We have above another interpretation of it as a method of testing the sample as a whole, based upon the criterion of likelihood, which will be valid as long as that criterion can be employed.

Whilst the *à priori* knowledge regarding the value of σ remains loosely defined, the open z-contours are probably the best that can be used. But if there are grounds for believing that σ must lie within a limited range, the contours which it would be natural to adopt become deflected until, when σ can be almost exactly located, they will form a closed system tending to the λ-contours. Suppose, for illustration, that in a sample of 10 there is a mean of 20·4 and a standard deviation of 3·6. We may ask whether it is likely that this sample has been drawn from a population with mean of 18·0. As long as there is no *à priori* information which enables the value of σ to be closely determined, or which is contradicted *à posteriori* by the sample value of $s = 3{\cdot}6$, it seems reasonable to apply Student's test. But if there were grounds for believing that σ could not exceed 1·5, the knowledge that for $z = \frac{2}{3}$, $P_z = {\cdot}26$ would be of little value in solving the problem in hand. This point is illustrated more fully in the first example of Section (9) below.

(6) *Alternative Method of examining Hypothesis B.*

Take as variables $M = \frac{m}{\sigma}$ and $S = \frac{s}{\sigma}$, and suppose Figure 3 (p. 192) to represent the (M, S)-plane. The λ contours of (xx) become

$$M^2 + S^2 - \log_e S^2 = k \log_e 10 \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(xxvi)},$$

and surround the point $C\,(M = 0, S = 1)$. Since the point-density in the (m, s)-plane is given by (v), that in the (M, S)-plane will be

$$d' = \text{constant} \times S^{n-2} e^{-\frac{n}{2}(M^2+S^2)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(xxvii)};$$

consequently this plane with density d', and the contour system (xxvi), can be looked upon as a fixed field of reference whatever be σ. In testing Hypothesis B, we cannot locate in the plane the point P corresponding to Σ, but only the radial line OT somewhere on which it must lie, where

$$M/S = m/s = z = \cot TOD.$$

P_z is the integral of the density throughout the sector TOD of the plane. If σ were known, the point P could be identified, falling let us suppose at Q. We could then find P_λ, the integral of d' over the region lying outside the λ-contour drawn through Q. P_λ will clearly be a maximum if P were found to fall at R, the point where OT touches a member of the λ-system. This has been seen from a different method of approach * to occur when $\sigma^2 = m^2 + s^2$, or $M^2 + S^2 = 1$. Consequently in Figure 3, $OR^2 = 1 = OC^2$.

We may now argue as follows. For a given Σ, the most favourable case giving a maximum value to P_λ arises when we suppose the sampled population from $M\,(\Pi)$ to be Π_1 with $\sigma = \sqrt{m^2 + s^2}$. If, on giving this value to σ, Hypothesis A judged by P_λ and the criterion of likelihood is improbable, then it will be more improbable for every other normal population with mean at a. Hence we reject Hypothesis B. If on the other hand P_λ is such that Hypothesis A appears reasonable for population Π_1, then Hypothesis B can hardly be rejected, because while it may not be possible to assert that it is most probably true, we know that for values of σ in the neighbourhood of $\sqrt{m^2 + s^2}$ the sample Σ is not exceptional. To the criticism that there is no justification for assuming that Π_1 is the population sampled, it can be answered that unless some *à priori* distribution of σ be postulated, there is no justification whatever for assuming that Π_1 or a population

* On p. 190 above. The result may also be proved directly as follows:
The contour equation is

$$\phi\,(M, S) \equiv M^2 + S^2 - \log_e S^2 = k \log_e 10,$$

and the tangent to this at the point (M_1, S_1) will be

$$S - S_1 = -(M - M_1) \frac{\partial \phi}{\partial M_1} \Big/ \frac{\partial \phi}{\partial S_1}.$$

But

$$\frac{\partial \phi}{\partial M_1} = 2M_1; \quad \frac{\partial \phi}{\partial S_1} = 2S_1 - \frac{2}{S_1},$$

so that the tangent becomes

$$S - S_1 = (M - M_1)\, M_1 S_1/(1 - S_1^2).$$

As this is to be the line OT passing through the origin,

$$S_1 = M_1^2 S_1/(1 - S_1^2) \text{ or } M_1^2 + S_1^2 = 1,\ m^2 + s^2 = \sigma^2.$$

differing slightly from it is improbable. The greater the range of values of σ above and below $\sqrt{m^2 + s^2}$, which give reasonable values to P_λ, the more probable becomes Hypothesis B, but no attempt is made to give this probability a unique numerical measure.

It will be seen from Figure 3 that if Student's test be applied, $2P_z$ is the integral of the field between the axis of M and the tangents OT, OT'. Clearly $2P_z < P_\lambda$ (max), and the one will be a single-valued function of the other. It follows that if we were to reject Hypothesis B whenever $P_\lambda(\max) \leqslant \alpha$, we should in fact always reject it for samples for which $z \geqslant \gamma$ and $2P_z \leqslant \beta$, where $\beta < \alpha$. If then a statistician thoughtlessly decides, whatever be the test, to reject an hypothesis when $P \leqslant \cdot 01$, say, and accept it when $P > \cdot 01$, it will make a considerable difference to his conclusions whether he uses P_z or P_λ (max). But as the ultimate

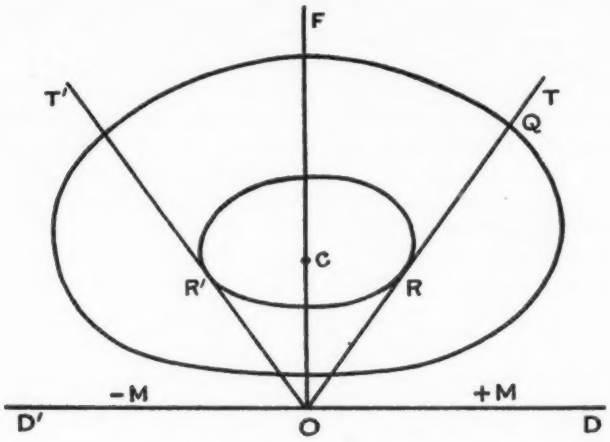


FIG. 3

value of statistical judgment depends upon a clear understanding of the meaning of the statistical tests applied, the difference between the values of the two P's should present no difficulty. The difference in the two scales of probability simply corresponds to the difference in attitude of mind with which the problem is approached; the use of P_λ (max) perhaps corresponds to the more cautious attitude, but it would seem impossible to claim that one approach is the correct one and the other is erroneous.

(7) *Solutions obtained by the Inverse Method.*

The inverse method of approach is to start from Σ, the observed sample, as the one certain fact in the problem. Points in the space no longer represent samples but populations from which Σ may have been drawn; that is to say, the density field is made appropriate to Σ and not to Π, and we shall reject an

hypothesis concerning the population sampled on entering regions of the field in which the density falls below a certain level. This alternative point of view is of some interest, and although it leads to tests which are almost identical with those reached above, we shall discuss it in some detail in the present case of sampling from a normal population.

We must again assume that Σ has been drawn from *some* normally distributed population, and as this can be described completely by its mean, a, and standard deviation, σ, the whole discussion may be confined to the two-dimensioned (a, σ)-plane. Referred to a fixed origin in this plane and rectangular axes, Σ is the point $G\ (\bar{X}, s)$ and any population Π a point $P\ (a, \sigma)$. The position is illustrated in Figure 4.

The ordinary method of inverse probability consists in postulating some function $\phi(a, \sigma)$ to represent the probability *à priori* that the sampled population is Π, and taking for the point-density of the field

$$D = \text{constant} \times \phi(a, \sigma) \times \sigma^{-n} e^{-\frac{n}{2\sigma^2}\{(\bar{X}-a)^2+s^2\}} \quad \text{.....(xxviii).}$$

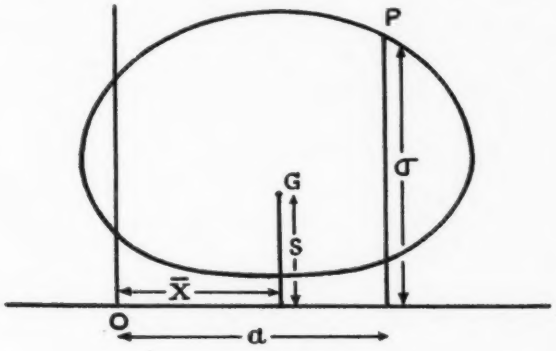


FIG. 4

$D.da.d\sigma$ would then be termed the probability *à posteriori* that the population sampled has a mean and standard deviation in the range, $a \pm \frac{1}{2}da$, $\sigma \pm \frac{1}{2}d\sigma$. The difficulty of this procedure in any practical problem lies in the fact that it is almost impossible to express ϕ in exact terms. We prefer therefore to follow a line of argument which while really equivalent to the above with ϕ assumed constant makes use of the principle of likelihood rather than the somewhat vaguer conception of *à posteriori* probability.

The likelihood of Π may be written

$$L \propto D = (1/\sqrt{2\pi})^n \sigma^{-n} e^{-\frac{n}{2\sigma^2}\{(\bar{X}-a)^2+s^2\}} \quad \text{.....(xxix).}$$

Population points such as P of Figure 4, for which L is constant will lie on closed curves,

$$\log_e \sigma^2 + \{(\bar{X}-a)^2 + s^2\}/\sigma^2 = h/n \quad \text{.....(xxx),}$$

which surround the sample point G. Give to the field a point-density D. If we are right in believing that the value of the likelihood is a useful criterion, then L (or D) itself would provide us with a measure of the confidence to be placed in an hypothesis as to Π. This quantity as it stands is not however measured on a very suitable scale, for it is difficult to know how to interpret the very small values that L will generally take. Let I be the integral of D taken throughout the whole (a, σ)-field*, and I' throughout that portion of it which lies outside the member of system (xxx) passing through P. Then the ratio I'/I appears to be a suitable measure to use. It is the same for all populations of a given likelihood and it varies between 0 and 1; when it is very small we know that the likelihood of Π stands very low in the scale, and when large that L approaches the maximum value. Likelihood as defined by Fisher is a quantity which cannot be integrated†. With this in the strict sense we agree, but look upon the relation $I'/I = f(L)$ as a form of transformation providing as a criterion a function of L placed on a scale more easily understood in common terms.

We can now write

$$P_{a,\sigma} = I'/I = c_0 \int_{\omega} \sigma^{-n} e^{-\frac{n}{2\sigma^2}\{(\bar{X}-a)^2+s^2\}} \, da\, d\sigma \quad \text{.....(xxxi),}$$

where the integral is taken outside the member of (xxx) passing through (a, σ), and c_0 is a quantity depending only on n and s, which is chosen so that $P_{a,\sigma} = 1$ when $a = \bar{X}$, $\sigma = s$.

It is of interest to compare the values of

$$P_{m,s}, \text{ (xii)}; \quad P_\lambda, \text{ (xxi)}; \quad \text{and } P_{a,\sigma}, \text{ (xxxi).}$$

Each is a function of n and the four quantities, $\bar{X}_0$, s_0, a_0 and σ_0, which define the particular Σ and Π with which Hypothesis A is concerned. Further, they are each integrals taken in a two-dimensioned density field over regions outside an oval contour. They may be brought into comparable form by means of the following transformations.

(*a*) $P_{m,s}$.

Transformation $\quad m/\sigma_0 = u, \quad s/\sigma_0 = v, \quad dm\,ds = \sigma_0^2\, du\, dv.$

Take $\quad k_0 = n\{\log(s_0^2/\sigma_0^2) - [(\bar{X}_0 - a_0)^2 + s_0^2]/\sigma_0^2\}.$

The density in the (u, v)-plane becomes, on transforming (v),

$$D' = c v^{n-2} e^{-\frac{n}{2}(u^2+v^2)} \quad \text{.....(xxxii).}$$

$P_{m,s}$ is obtained by integrating D' outside the contour

$$(n-2)\log v^2 - n(u^2+v^2) = k_0 - 2\log(s_0^2/\sigma_0^2) \quad \text{.....(xxxii } bis\text{).}$$

* It is not necessary to suppose this to range from $a = +\infty$ to $-\infty$, $s = 0$ to $+\infty$. The conclusions will not be modified if we suppose that the field is bounded by one of the contours of (xxx) which contains, let us say, ·999 of the density field.

† See *Phil. Trans.* A, Vol. 222, p. 327.

The integral I does not equal unity, but $\Gamma\left(\frac{n-2}{2}\right) \Big/ \{2^{\frac{n}{2}} \Pi^{\frac{n-1}{2}} n^{\frac{n-1}{2}} s^{n-2}\}$.

(*b*) P_λ.

The transformation is as for $P_{m,s}$, and the density in the (u, v)-plane is as in (xxxii). The integral is to be taken outside the contour

$$n \log v^2 - n(u^2 + v^2) = k_0 \quad \text{.....(xxxiii).}$$

This corresponds to (xxvi).

(*c*) $P_{a,\sigma}$.

The transformation is $\quad (\bar{X}_0 - a)/\sigma = u, \ s_0/\sigma = v,$

whence $\quad a = X_0 - s_0 u/v, \ \sigma = s_0/v,$

and $\quad da\,d\sigma = J \,.\, du\,dv = s_0^2/v^3\, du\,dv.$

Then the density becomes, using (xxix),

$$D'' = c' v^{n-3} e^{-\frac{n}{2}(u^2+v^2)} \quad \text{.....(xxxiv).}$$

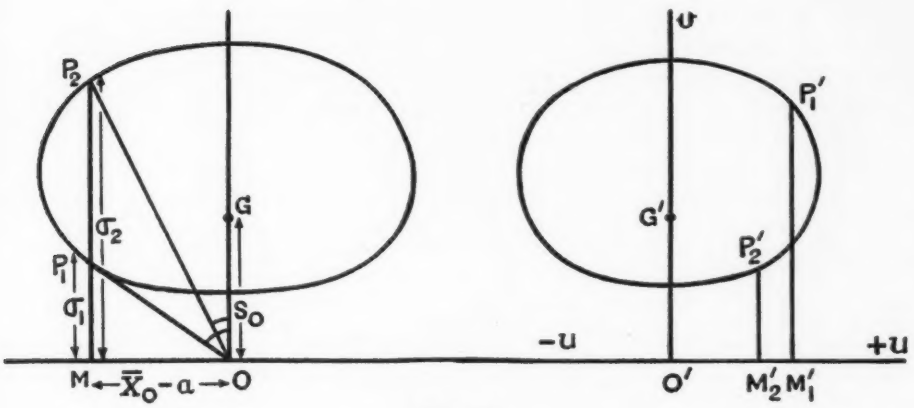


FIG. 5

And the contour, using (xxx), is

$$\log(s_0^2/v^2) + u^2 + v^2 = \log \sigma_0^2 + (\bar{X}_0 - a_0)^2/\sigma_0^2 + s_0^2/\sigma_0^2,$$

or $\quad n \log v^2 - n(u^2 + v^2) = k_0 \quad \text{.....(xxxiii } bis\text{).}$

The nature of this transformation is shown in Figure 5. The sample point G transforms into G', $u = 0$, $v = 1$, and two points P_1, P_2 on a contour become P_1', P_2', where

$$O'M_1' = u_1 = \tan GOP_1, \quad P_1'M_1' = v_1 = OG/P_1M,$$
$$O'M_2' = u_2 = \tan GOP_2, \quad P_2'M_2' = v_2 = OG/P_2M.$$

The field outside the contour in the (a, σ)-plane corresponds to that outside the contour in the (u, v)-plane and it is over this that D'' is to be integrated.

On comparing these results in the (u, v)-plane we see that

(1) the density fields for $P_{m,s}$ and P_λ are the same, namely that given by (xxxii), but the contours (xxxii *bis*) and (xxxiii) differ both in the factor $(n-2)$ for n, and in the additional constant term $-2\log(s_0^2/\sigma_0^2)$ contained in the former;

(2) the density fields for P_λ and $P_{a,\sigma}$ differ in the term v^{n-2} of (xxxii) and v^{n-3} of (xxxiv), but the contours are identical (xxxiii).

In large samples, for a given Σ and Π, the three tests become practically equivalent since the values of the three P's will scarcely differ. They each result however from different methods of approaching the problem, and in small samples will lead to somewhat different results. It is the quantity P_λ which has been tabled.

We shall now consider whether it is possible to obtain from the (a, σ)-field, measures of the probability that the population, from which Σ has been drawn, has

(1) a mean a, such that $a_1 \leqslant a \leqslant a_2$, irrespective of the standard deviation;

(2) a standard deviation σ, such that $\sigma_1 \leqslant \sigma \leqslant \sigma_2$, irrespective of the value of a.

Again, as likelihood can only be associated with a single population with unique values of a and σ, we cannot speak of the likelihood of populations for which (a, σ) falls within the regions of the field defined above, although we might speak of the probability *à posteriori* of such an event, assuming $\phi(a, \sigma)$ to be constant. Whatever be its exact interpretation, however, the integral of D in (xxix) taken over these regions and divided by I (the integral over the whole field) will provide us with some form of measure of the probability of the hypotheses we are considering, and will be a quantity varying between 0 and 1.

First integrate D, therefore, along a strip of breadth da at a distance a from the axis of σ:

$$da\int_0^\infty D\,.\,d\sigma = \text{constant} \times da\int_0^\infty \sigma^{-n}\, e^{-\frac{n}{2\sigma^2}\{(\bar{X}-a)^2+s^2\}}\, d\sigma.$$

Write $u^2 = n\{(\bar{X}-a)^2+s^2\}$ and $u^2/2\sigma^2 = v$, so that $-d\sigma\, u^2/\sigma^3 = dv$ and the integral becomes

$$= \text{constant} \times u^{-n+1}\, da\int_0^\infty e^{-v} v^{\frac{n-3}{2}}\, dv$$

$$= \text{constant} \times u^{-n+1}\, da,$$

the constant being a function of n only.

Dividing by I, and writing $z = (\bar{X}-a)/s$, we have finally for the measure of probability that the mean of Π is such that $a_1 \leqslant a \leqslant a_2$,

$$P(a_1, a_2) = c_7\int_{z_1}^{z_2}(1+z^2)^{-\frac{n-1}{2}}\, dz \quad \text{.....(xxxv)},$$

where $z_1 = (\bar{X}-a_1)/s, \quad z_2 = (\bar{X}-a_2)/s.$

Except that we have now $n-1$ instead of n this expression corresponds to (xi), and its value may be obtained from the tables of Student's z or t function, using $n-1$ instead of n*.

Now integrate D along a strip of breadth $d\sigma$ at a distance σ from the axis of a:

$$d\sigma \int_{-\infty}^{+\infty} D \, . \, da = \text{constant} \times \sigma^{-n+1} e^{-\frac{ns^2}{2\sigma^2}} d\sigma.$$

On dividing by I we have for the measure of probability that the standard deviation of Π is such that $\sigma_1 \leqslant \sigma \leqslant \sigma_2$,

$$P(\sigma_1, \sigma_2) = c_s \int_{\sigma_1}^{\sigma_2} \sigma^{-n+1} e^{-\frac{ns^2}{2\sigma^2}} d\sigma \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(xxxvi)}.$$

By making the transformation $\sqrt{n}s/\sigma = \psi$, this becomes

$$P(\psi_1, \psi_2) = \int_{\psi_1}^{\psi_2} \psi^{n-2} e^{-\frac{1}{2}\psi^2} d\psi \Big/ \int_0^{\infty} \psi^{n-2} e^{-\frac{1}{2}\psi^2} d\psi \quad \ldots\text{(xxxvi } \textit{bis}\text{)},$$

which is the difference of two ordinary χ^2 integrals and can be obtained from the χ^2 Tables by taking $\psi^2 = \chi^2$, $n' = n-1$.

If in equation (ix), where the variable is s, we write $\psi = \sqrt{n}s/\sigma$, we obtain a probability integral P_s of exactly the same type as (xxxvi), except that there will be a term ψ^{n-2} instead of ψ^{n-3}, that is to say, the χ^2 Tables must be entered with $n' = n$.

We have reached, then, three pairs of corresponding results, (xxi) and (xxxi), (xi) and (xxxv), (ix) and (xxxvi); in each case there is this difference of a unit in n which appears to result invariably on proceeding by the inverse method. Probably most statisticians will prefer to use the criteria (xxi), (xi) and (ix) of the direct method, which with the addition of the present P_λ Tables can be readily expressed in numerical terms. The difference of the two is however of no great importance except in very small samples, where final conclusions will be drawn in any case with some hesitation.

(8) *Analysis of Church's Samples from a Skew Population.*

Dr A. E. R. Church very kindly placed at our disposal one of his sets of sampling results previously obtained for another purpose†. These were the 1000 samples of 10 drawn from a smooth distribution with the following frequency constants:

$$\sigma = 3{\cdot}569{,}207, \quad \beta_1 = {\cdot}219{,}333, \quad \beta_2 = 3{\cdot}157{,}676.$$

This distribution not differing very greatly from a Type III form may be taken as typical of the moderately skew distributions frequently met with in

* Student's original table of $z = m/s$ for $n = 2$ to 10 is given in *Tables for Statisticians and Biometricians*, p. 36 (second edition). A complete table for $n = 2$ to 21 with auxiliary tables for use with higher values has been published in *Metron*, Vol. v. p. 3, 1925. These latter tables must be entered with $t = z\sqrt{n'-1}$, where n' is now taken as the size of sample.

† *Biometrika*, Vol. XVIII. pp. 321—324. This is Church's Population "B" and group of samples "IV."

practice. It is therefore of considerable importance to discover how far the sampling variation in the frequency constants can be represented by the appropriate "normal" formulae. In practice we may often not have available sufficient information to test whether the population II be normally distributed or no, and even if this be available may be unable to deduce the appropriate formulae for

TABLE I.

Distribution of Frequency Constants for Church's 1000 *Samples of* 10.

Frequencies for S			Frequencies for z		
Central values of S	Observed	Normal Theory	Central values of z	Observed	Normal Theory
			Less than $-1{\cdot}20$	—	2·9
—	—	—	$-1{\cdot}15$	2	1·7
—	—	—	$-1{\cdot}05$	3	2·9
1·3	1·5	20·3	$-\ {\cdot}95$	3	4·7
1·5	0·5		$-\ {\cdot}85$	8	7·7
1·7	11		$-\ {\cdot}75$	10	12·6
1·9	21	21·2	$-\ {\cdot}65$	6	20·2
2·1	32·5	34·3	$-\ {\cdot}55$	38	31·2
2·3	71	50·3	$-\ {\cdot}45$	46	46·4
2·5	68	67·0	$-\ {\cdot}35$	59	65·5
2·7	81	82·5	$-\ {\cdot}25$	92	85·7
2·9	81·5	94·0	$-\ {\cdot}15$	119	103·9
3·1	101·5	100·1	$-\ {\cdot}05$	134	114·6
3·3	89·5	99·8	$+\ {\cdot}05$	103	114·6
3·5	80	93·9	$+\ {\cdot}15$	104	103·9
3·7	95	83·2	$+\ {\cdot}25$	78	85·7
3·9	67	69·8	$+\ {\cdot}35$	63	65·5
4·1	54	55·7	$+\ {\cdot}45$	36	46·4
4·3	53	42·2	$+\ {\cdot}55$	39	31·2
4·5	33	30·5	$+\ {\cdot}65$	21	20·2
4·7	19·5	21·0	$+\ {\cdot}75$	9	12·6
4·9	11·5	13·9	$+\ {\cdot}85$	11	7·7
5·1	9	8·7	$+\ {\cdot}95$	6	4·7
5·3	6	5·2	$+1{\cdot}05$	4	2·9
5·5	6	6·4	$+1{\cdot}15$	2	1·7
5·7	4		$+1{\cdot}25$	—	1·1
5·9	2		$+1{\cdot}35$	1	0·7
6·1	—		$+1{\cdot}45$	1	0·4
6·3	1		$+1{\cdot}55$	1	0·2
—	—	—	$+2{\cdot}05$	1	0·5
—	—	—	Greater than $+2{\cdot}10$	—	
Total	1000·0	1000·0	Total	1000	1000·0

Frequency Constants for s			
	Observed	Normal Theory	Standard Error
Mean	3·3181	3·2933	·0249
Standard Deviation	·8086	·7863	·0176 (approx.)
β_1	·1467	·0634	—
$\sqrt{\beta_1}$	+·3830	+·2518	·0761
β_2	2·9782	3·0106	·176

Frequency Constants for z			
	Observed	Normal Theory	Standard Error
Mean	+·0120	0	·0120
Standard Deviation	·3709	·3780	·0107
β_1	·1740	0	—
$\sqrt{\beta_1}$	+·4171	0	·181
β_2	4·5405	4·2000	1·363

Frequencies outside P_λ Contours in the (m, s)-field

k of Contour	Observed	Normal Theory
·45	847	849
·50	500	504
·55	321	299
·60	181	177
·65	102	105
·70	59	63
·80	19	22
·90	8	8

samples from skew distributions. To what extent, then, shall we be likely to draw faulty conclusions if we apply the normal criteria in these cases?

Church has fitted a normal curve with standard deviation of $\sigma/\sqrt{n}$ to the distribution of means, and finds on applying the test for goodness of fit, a value for P of ·5758 *. By taking the square roots of his values of s^2, we have obtained the distribution of s given in Table I. This is compared in the next column with the theoretical distribution appropriate to samples from a normal distribution with the same standard deviation, viz. $\sigma = 3{\cdot}569{,}207$. The distribution of s^2 for a normal population is of course a Type III curve, and the theoretical frequencies for s were found by taking the areas between the corresponding ordinates of the s^2 curve from the Tables of the Incomplete Gamma Function. Testing for goodness of fit, $\chi^2 = 26{\cdot}12$ and with $n' = 19$, $P = {\cdot}097$. The larger of the differences in frequency are rather irregularly scattered, but there appear on the whole to be too few low and too many high values of s among the observations. This means that the observational distribution is more skew than that of normal theory, a point brought out from a comparison of the frequency constants given in the same table. The constants for normal theory are taken from the table given in *Biometrika*, Vol. x. p. 529. The standard error for σ_s was obtained on the assumption that the theoretical distribution of s was normal†. The ratios of the differences between the observed and theoretical values of the four frequency constants to their standard errors are $+1{\cdot}0$, $+1{\cdot}2$, $+1{\cdot}69$ and $-0{\cdot}2$ respectively; the disagreement except perhaps in the case of $\sqrt{\beta_1}$ cannot be considered very serious, although we are probably just reaching a degree of skewness in the population for which the normal theory begins to fail.

Compare now the observed distribution and that of normal theory for $z = m/s$. They are given in Table I; here $\chi^2 = 29{\cdot}19$, $n' = 18$ and $P = {\cdot}033$. The most serious discrepancy, contributing 9·98 to χ^2, occurs in the group for $z = -{\cdot}6$ to $-{\cdot}7$, and there is no doubt that among the observations there are too few with large negative and too many with large positive values of z. The observed z-distribution is in fact significantly skew, whereas Student's distribution is symmetrical. A comparison of the frequency constants is given in the same table. Judged by the standard errors, it is only $\sqrt{\beta_1}$ that differs significantly from normal theory‡.

* *Biometrika*, Vol. XVIII. p. 334.

† For low values of β_1 the table (*Tables for Statisticians and Biometricians*, Table XXXVII) giving the standard error of this quantity cannot be used, as in this region the second order term is as great or greater than the first order term which vanishes when $\beta_1 = 0$. It is the first order term which is tabled. We have therefore compared the values of $\sqrt{\beta_1}$ for which the standard error to first order is equal to

$$\{\beta_4 - 3\beta_3 - 6\beta_2 + \tfrac{9}{4}\beta_1\beta_2 + \tfrac{3}{4}\beta_1 + 9\}^{\frac{1}{2}}/\sqrt{N}$$

(where N is number of samples), and have inserted values for β_3 and β_4 obtained from the theoretical β_1 and β_2 on the assumption that the distribution can be represented by a Pearson curve.

‡ The standard error of $\sqrt{\beta_1}$ was calculated from the formula given in the preceding footnote; this gives to the first order for samples from symmetrical distributions,

$$(\text{S.E. } \sqrt{\beta_1})^2 = (\beta_4 - 6\beta_2 + 9)/N.$$

In the case of Student's distribution for $n = 10$, $\beta_4 = 49{\cdot}0$.

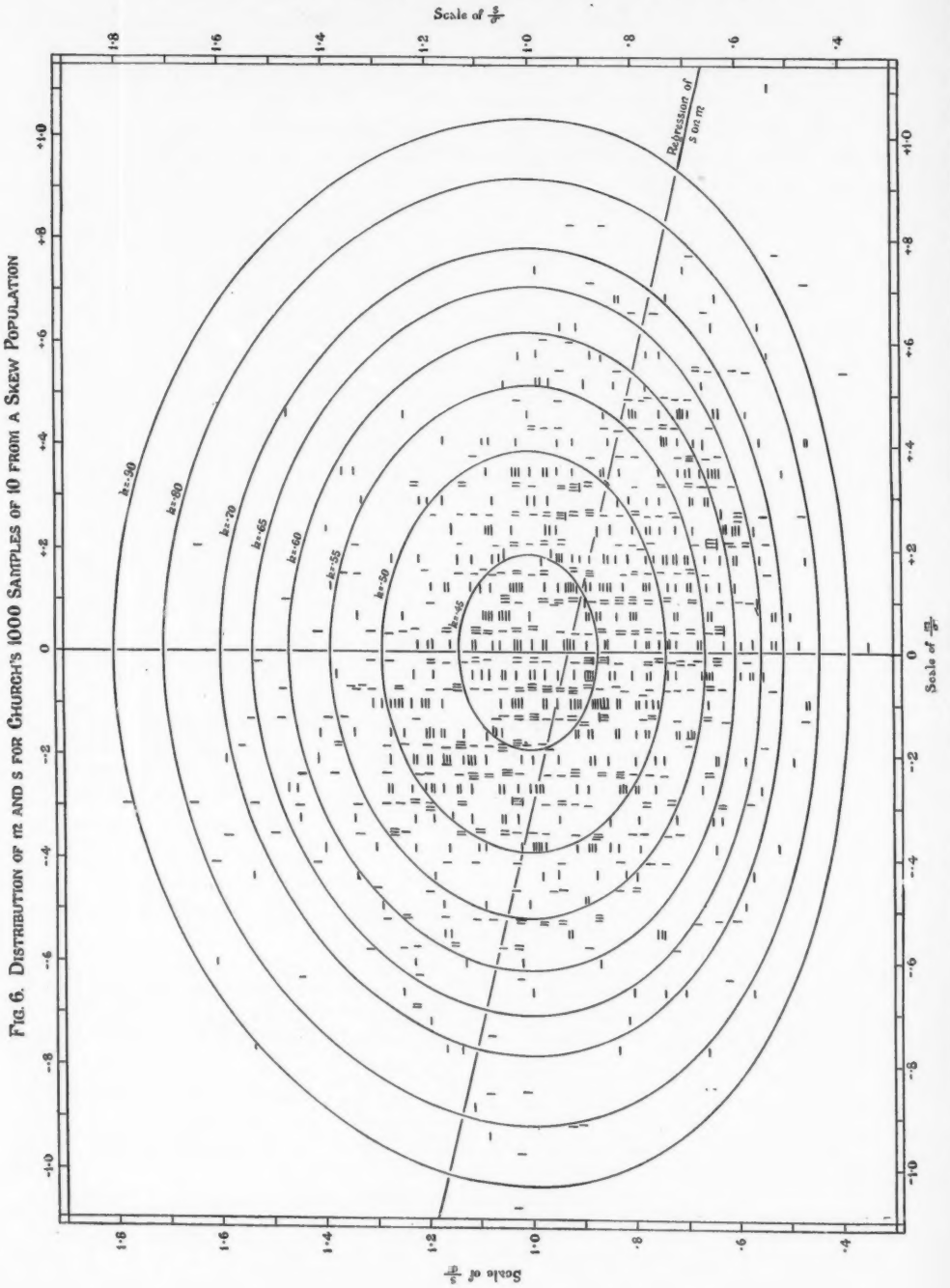


FIG. 6. DISTRIBUTION OF m AND s FOR CHURCH'S 1000 SAMPLES OF 10 FROM A SKEW POPULATION

Church's population distribution was negatively skew, and this results in positive skewness in the distribution of z.

In practice the method of applying the test consists, when the observed value of z is large, in rejecting Hypothesis B with a degree of confidence which increases as P_z decreases. Taking P_z equal (to the nearest figure) to (*a*) ·05, and (*b*) ·005, we find from the observations,

(*a*) { for positive values of z: 57 cases out of 1000 instead of 53,
" negative " " : 32 " " " 53,

(*b*) { " positive " " : 6 " " " 5,
" negative " " : 2 " " " 5.

This difference must not be passed over, but it is doubtful whether the knowledge that P_z was really ·03 (or ·06) rather than ·05, or again ·002 (or ·006) rather than ·005, would in fact ever modify our judgment when balancing the probabilities regarding the origin of a single sample.

Finally we may consider the distribution of the observations among the P_λ contours. Figure 6 is a reproduction of a rough working diagram containing entries for the 1000 doublets (m, s). The scale attached to the margins is that of the ratios m/σ and s/σ, where σ is as before the standard deviation of the population. Eight of the P_λ contours are drawn in the figure; the observed regression straight line of s on m has also been given. It is obvious from inspection that the condition which holds for samples from a normal distribution of no correlation between m and s and of homoscedasticity is now far from true. The coefficient of correlation calculated from the observations is $r = +\cdot3206$ with standard error ·0191. The theoretical correlation between m and s^2, which gives a fair approximation to that between m and s, is ·3036*. The correlation ratio of s on m is $\eta = \cdot3495$. In spite of this obvious disagreement with the distribution of the normal (m, s)-field, it is found that the frequencies of sample points lying outside the successive P_λ contours follow very closely the expected numbers. A comparison is shown at the end of Table I; the only disagreement of importance is that for the contour, $k = \cdot55$, outside which 321 observations instead of 299 lie. The χ^2-test may be applied to the frequencies *between* contours, for which there will be 9 groups including that for observations inside $k = \cdot45$ and for observations outside $k = \cdot90$; here $\chi^2 = 7\cdot36$ and $P = \cdot50$.

The P_λ contours can no longer of course provide a completely satisfactory criterion, since they have ceased to divide on a logical basis the sample points for which Hypothesis A is more probable from those for which it is less probable. They appear however to enable us to control the first of the two sources of error described in the introductory section†, since they do enclose approximately the expected proportions of random samples drawn from Π. Whether the situation would be as satisfactory for samples of a different size, or for different population distributions, we cannot of course say without further experiment.

* Calculated from the formula (68) given in *Biometrika*, Vol. XVII. p. 479.

† See p. 177 above.

(9) *Illustrations of the use of the Tables of* P_λ.

The tables are given in the Appendix with a description of the method by which they were computed and of how they may be extended. They must be entered with n, the size of sample, and k; in the following examples k has been obtained from one or other of the diagrams, Figures 13 and 14, pp. 236, 237, in which the marginal scales are (1) M, the ratio of the deviation in the sample mean from the population mean to the population standard deviation σ, and (2) S, the ratio of the sample standard deviation, s, to σ.

Example 1.

Measurements on 884 Egyptian skulls (XXVI—XXX Dynasties)* gave for the distribution of cephalic index, Mean = 75·06, Standard Deviation = 2·68. Would it be justifiable to consider that the 10 skulls with cephalic indices as follows, 74·1, 77·7, 74·4, 74·0, 73·8, 72·2, 75·2, 78·2, 77·1, 78·4, were a random sample from this population?

(*a*) The distribution of cephalic indices is in general found to be symmetrical and leptokurtic; we shall assume, pending the completion of further inquiries, that if $\beta_2 = 3{\cdot}4$ to $3{\cdot}5$, the P_λ contours are still adequate. It is found that the mean cephalic index is 75·51 and $s = 2{\cdot}059$, from which $M = +\,{\cdot}168$ and $S = {\cdot}768$. Using Figure 13 (the sign of M is immaterial) it is seen that the sample point lies just inside the contour $k = {\cdot}50$; entering the P_λ Table of p (see p. 238) with $n = 10$ and $k = {\cdot}50$ it is found that $P_\lambda = {\cdot}504$. It is therefore possible to say that *judged by cephalic index alone* there would be no reason to doubt that these 10 skulls were a random sample of XXVI—XXX Dynasty material†.

(*b*) Now apply the same test to the following 10 cephalic indices: 66·7, 69·4, 67·8, 73·2, 79·3, 80·7, 64·9, 82·2, 72·4, 78·1. Here $\overline{X} = 73{\cdot}47$, $s = 5{\cdot}942$, $M = -\,{\cdot}593$, $S = 2{\cdot}217$; the sample point falls near the contour $k = 1{\cdot}60$ in Figure 14, and for $n = 10$ this corresponds to P_λ less than ·0001. From the position of the point in the diagram it will be seen that the divergence is mainly due to a very large value of s compared with σ; the mean is not exceptionally divergent from the supposed population mean of 75·06. We should conclude that it was very improbable that the 10 skulls were a random sample from the Egyptian population, and that probably they consisted of heterogeneous material, the variability being greater than that generally found in a sample from a homogeneous race.

(*c*) Suppose that we had been given only the mean cephalic index and not the standard deviation of the population. We could ask whether it was likely having regard to the observed variability of $s = 5{\cdot}942$, that a random sample of 10 could have been drawn having a mean differing by 1·59 from the population mean. Applying Student's test‡ it is found that $z = -\,{\cdot}268$ or $t = z\sqrt{n-1} = -\,{\cdot}803$

* *Biometrika*, Vol. XVI. Tables II and III, pp. 337, 338.

† Actually the cephalic indices are those of the first 10 of the Farringdon Street skulls measured by Miss Hooke. *Biometrika*, Vol. XVIII. p. 54, Appendix A.

‡ Using either of the Tables referred to in the footnote to p. 197 above.

and $P_t = P_z = \cdot 221$. That is to say so large a deviation of z in excess or defect from zero would occur in 44 per cent. of random samples. Consequently the evidence does not contradict Hypothesis B*.

(*d*) If however we had known from other sources that the standard deviation of cephalic index within a single race lay within the range 2·5 to 4·0, we could have improved upon the z-test†. Although σ be unknown, we could say that the sample point in Figure 14 must lie somewhere on the line $z = \cdot 268$ which joins the origin, $M = 0 = S$, to the corresponding division of the radiating z-scale given in the right-hand margin of the diagram. The exact position will depend on the value given to σ, but if we might infer that this is unlikely to be greater than 4·0, we could say that it was very unlikely that the sample point lay farther to the left along the z-line than the point for which $S = s/\sigma = 5 \cdot 942/4 = 1 \cdot 485$. To the right of this point the contours cutting the line have $k > \cdot 68$ and therefore $P_\lambda < \cdot 08$. We should therefore argue that it was far less likely that the hypothesis as to mean cephalic index of population were true than the value of P_z obtained in (*c*) would suggest.

(*e*) If there had been no *à priori* information as to σ, it would still have been possible to locate the sample point on the line $z = \cdot 268$. By taking $\sigma = \sqrt{m^2 + s^2} = 6 \cdot 13$, or $M^2 + S^2 = 1$, we should locate the sample at that point on the line where it touches a contour of the system. This will roughly be that with $k = \cdot 47$, giving $P_\lambda = \cdot 689$. Consequently were it quite possible for a homogeneous race to have a standard deviation of cephalic index as great as 6·0, we should have no reason for doubting the origin of the second sample of 10. This is an illustration of the line of reasoning followed in Section (6) above.

Example 2.

Records of weight in a large population of mice for males between 120 and 140 days of age show a mean of 23·823 grammes, $\sigma = 3 \cdot 137$ grammes, and for the frequency distribution $\beta_1 = \cdot 086$, $\beta_2 = 2 \cdot 687$.

Can the groups of 6, 4 and 5 mice with the following weights be considered random samples from this population?

(*a*) 22·5, 26·0, 20·5, 24·0, 18·0, 24·5.

(*b*) 21·0, 21·0, 20·0, 20·5.

(*c*) 22·5, 23·5, 23·5, 25·0, 24·5.

(*a*) $n = 6$, $M = -\cdot 395$, $S = \cdot 851$. Using Figure 13, we see that $k = \cdot 52$ approximately, giving $P_\lambda = \cdot 607$, so that there is no reason to doubt the origin of the sample.

* If the inverse argument discussed in Section (7) above is used, Student's Tables would be entered with $n = 9$, and this gives $P_z = \cdot 223$ against $\cdot 221$, a difference of no importance as far as interpretation is concerned.

† The following values for the standard deviation of cephalic index will be found in recent papers in this Journal: XXVI—XXX Dynasty Egyptian 2·68; Naqada 2·80; 17th Century English 3·26; Farringdon Street 3·48; Australian "A" 3·88.

(*b*) $n = 4$, $M = -1{\cdot}020$, $S = {\cdot}132$. The point is beyond the outermost contour of the diagram but this suggests a value for k of 2·1. Using the exact equation for k, or

$$k = {\cdot}434{,}294\,(M^2 + S^2) - 2\log_{10} S,$$

it is found that $k = 2{\cdot}218$ and $P_\lambda = {\cdot}002$. This small value of P_λ is chiefly due to the low value of S.

(*c*) $n = 5$, $M = -{\cdot}007$, $S = {\cdot}278$. Here $k = 1{\cdot}15$, $P_\lambda = {\cdot}036$. Again the value of S is unusually small.

The values of P_λ in (*b*) and (*c*) are small enough to make us search for an alternative hypothesis to that of pure chance fluctuation. As a matter of fact while (*a*) was a randomly chosen sample, in both the cases (*b*) and (*c*) the mice belonged to a single litter, so that the individual weights were correlated partly from the effect of inheritance, and partly because weight is correlated with size of litter.

Example 3.

The tables may be used in cases where the normal curve gives an approximation to the binomial.

A count of over a thousand words taken at random from daily newspapers gave a proportion of ·6691 monosyllables. Out of 35 words chosen at random from the works of 12 English authors, the following numbers of monosyllables were found*: Kipling 26, Thackeray 26, Macaulay 24, Meredith 24, Carlyle 23, Scott 23, Stevenson 22, Conrad 22, Borrow 22, Ruskin 21, Bryce 19, Hazlitt 12. Do these representatives of literature either appear to use a lower proportion of monosyllables or show among themselves a greater diversity in proportion than would be expected to occur in random samples from a vocabulary in which the proportion of monosyllables was ·6691?

The appropriate binomial has $n = 12$, $p = {\cdot}6691$, $q = {\cdot}3309$, and therefore a mean of 23·42 and standard deviation of 2·784 monosyllables. The $\beta_1 = {\cdot}015$ and $\beta_2 = 2{\cdot}957$, so that the fit by a normal curve should be satisfactory. The 12 samples of 35 have a mean of 22·00 and standard deviation of 3·559, giving $M = -{\cdot}510$, $S = 1{\cdot}278$, and $k = {\cdot}60$, $P_\lambda = {\cdot}121$. The P_λ would be very much larger but for the sample from Hazlitt†, and there appears in this particular character to be no significant difference between the language of the 12 writers and that of the newspapers.

Example 4.

Another type of problem may be illustrated as follows. A number of laboratories are perhaps determining by chemical analysis the percentage of some constituent substance in a given material. Take as an example the hypothetical results

* The 35 words were the first on the last line of each of 35 consecutive pages.

† A further random sample of 100 words taken from Hazlitt's *Characters of Shakespear's Plays* contained 42 monosyllables, corresponding to 14·7 out of 35, and suggesting that there may really be some significant difference in Hazlitt's vocabulary!

given in the following table; for each laboratory there is a mean per cent. m_t based on n_t analyses; the weighted mean of the whole is 8·353 per cent. and the combined standard deviation of the 187 analyses ·329 per cent. Information is

TABLE II.

Laboratory	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Mean per cent., m_t	8·432	8·127	8·483	8·002	8·317	8·582	8·239	8·531	8·163	8·405	8·391	7·976
No. of analyses, n_t	22	10	19	15	16	25	9	16	4	16	23	12
$(m_t - 8·4)\sqrt{n_t}/\sigma$	+·46	−2·63	+1·10	−4·69	−1·01	+2·77	−1·47	+1·59	−1·44	+·06	−·13	−4·47

required as to whether, (*a*) the results are consistent with a quoted trade percentage of 8·4 for the constituent substance, (*b*) the variation in means is no more than would be expected to arise from chance fluctuations, or (*c*) that variation points to systematic differences in the method of analysis or in the material distributed among the different laboratories. If the variation among the 187 observations is simply due to chance errors, then $\sigma = ·329$ may be taken as a fairly close measure of the standard deviation of the universe of possible variations in analysis which has been sampled; we may therefore test the hypothesis that the sampled population mean is 8·4 and that the grouping of the observations into 12 sets is of a purely random nature, in the following manner. The standard error of each laboratory mean will be $\sigma/\sqrt{n_t}$, and, if the hypothesis be true, the 12 quantities $(m_t - 8·4)\sqrt{n_t}/\sigma$ given in the last row of the table should each be a chance fluctuation from zero reduced to a standard error of unity. On the assumption, which could be verified in a given case, that the distribution of individual analyses is not exceptionally anormal, the distribution of these deviations in mean may be taken as normal. Consequently if we calculate the mean and standard deviation of the figures in the last row of the table, we can test the hypothesis by taking these as M and S, and finding k and P_λ with $n = 12$.

It is found that $M = -·821$, $S = 2·192$, $k = 1·70$, and P_λ is less than ·0001. The hypothesis is therefore extremely improbable and from the position of the point on Figure 14 it is clear that this is mainly due to the high value of S. Comparing the difference between the weighted mean of the 187 observations (8·353), and 8·4 with its standard error of $\sigma/\sqrt{187} = ·0240$, we find a ratio of −1·94, and a deviation as great as this in excess or defect might be expected to occur in 5 per cent. of random samples. The variation among the means themselves is however quite inconsistent with the hypothesis of chance fluctuations from a general mean, and must probably be attributed to variation in technique or to lack of homogeneity in the material distributed for analysis.

Example 5.

It is of some interest to compare different methods of approaching the problem of determining whether two samples can have been drawn from the same population. Suppose that the following figures were the results of an experiment

in Industrial Psychology. A piece of work has been carried out by one group of 30 workmen according to Method I, and by another group of 40 according to Method II, the groups being selected at random from a factory. Are the recorded distributions in times such as to justify the conclusion that Method I is the speedier of the two?

TABLE III.

Time in seconds		50	51	52	53	54	55	56	57	58	59	60	Totals
Frequency Distributions	Method I	1	3	5	4	7	5	3	1	1	—	—	30
	Method II	—	1	2	5	8	9	6	3	3	1	2	40

It is found on calculation that,

for I, mean, $\bar{x}_1 = 53{\cdot}700$ secs.; standard deviation, $s_1 = 1{\cdot}882$ secs.

for II, „ $\bar{x}_2 = 55{\cdot}175$ secs.; „ „ $s_2 = 2{\cdot}072$ secs.

The usual method of procedure is to compare the difference in means with the standard error of that difference, and to refer the ratio of the two to the appropriate probability tables. The value given to the standard error and the tables used will depend entirely on the line of reasoning employed in the final process of inference.

(*a*) We may compare $\bar{x}_1 - \bar{x}_2$ with $\sqrt{s_1^2/n_1 + s_2^2/n_2}$, and from the ratio, $-3{\cdot}107$, obtain $\frac{1}{2}(1+\alpha) = {\cdot}9991$ from Sheppard's Tables, and from this conclude that if the samples came from two populations with the same mean but standard deviations of s_1 and s_2, so great a difference in mean, positive or negative, would only occur in about two samples in 1000.

(*b*) Having regard to the mean values of s_1 and s_2 in sampling we may attempt to obtain estimates of σ_1 and σ_2 which in the long run are likely to be closer than s_1 and s_2, namely: $s_1\sqrt{n_1/(n_1-1)}$ and $s_2\sqrt{n_2/(n_2-1)}$, substitute these in the standard error, and hence obtain a value of the ratio $= -3{\cdot}060$ and a resulting $\frac{1}{2}(1+\alpha) = {\cdot}9989$.

(*c*) We may give to the result, (*b*), the inverse interpretation suggested by Rhodes*.

(*d*) We may approach the problem by making use of the principle of likelihood and reaching the test given by Fisher†. Suppose that we have reason to believe that two samples have been drawn from normal populations with the same standard deviation σ, but that it is necessary to compare the relative probability of two hypotheses, (1) that the samples come from identical populations with mean at a, and (2) that while identical as to variability there are two different means a_1 and a_2.

* *Journal of Royal Statistical Society*, 1926, Vol. LXXXIX. p. 544.

† *Metron*, 1925, Vol. v. p. 7. *Statistical Methods for Research Workers*, 1925, p. 109.

Then it is easy to show that the maximum likelihood of hypothesis (1) is associated with a population for which

$$a = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2};\ \sigma^2 = \frac{1}{n_1 + n_2}\left\{n_1 s_1^2 + n_2 s_2^2 + \frac{n_1 n_2}{n_1 + n_2}(\bar{x}_1 - \bar{x}_2)^2\right\},$$

giving

$$L\,(\max) \propto \frac{1}{\sigma^{n_1+n_2}} e^{-\frac{n_1+n_2}{2}}.$$

The maximum likelihood of hypothesis (2) is associated with a population for which

$$a_1 = \bar{x}_1;\ a_2 = \bar{x}_2;\ \sigma'^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2},$$

and

$$L'\,(\max) \propto \frac{1}{\sigma'^{n_1+n_2}} e^{-\frac{n_1+n_2}{2}}.$$

As a measure of comparison of the two hypotheses we may take the ratio of these two maximum likelihoods, or

$$L\,(\max)/L'\,(\max) = (\sigma'/\sigma)^{n_1+n_2} = (1 + z^2)^{-\frac{n_1+n_2}{2}},$$

where

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{n_1 s_1^2 + n_2 s_2^2}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}.$$

The smaller be z, the less likely becomes the hypothesis of a single population compared to the alternative hypothesis of two populations differing in mean, and consequently we can take z or a function of z as a criterion to use in judging the probability that the two samples have come from the same population. It is most convenient to refer z to the sampling distribution that it would follow were hypothesis (1) true. This Fisher has shown to be

$$y = y_0 (1 + z^2)^{-\frac{n_1+n_2-1}{2}},$$

so that in using the test we have only to enter Student's Tables with

$$t = \sqrt{n_1 + n_2 - 2}\,.\,z$$

and $n' = n_1 + n_2 - 1$ *.

In the present example $z = -\cdot3662$, $t = -3\cdot020$ and $P_z = P_t = \cdot9982$.

This extension by Fisher of Student's test would seem to be of considerable value in dealing with small samples. Its limitation lies partly in the assumption of normality, but as the z-distribution is not very sensitive to changes in the population β's this will often not be serious; and partly in the fact that it does not take into account a possible difference in population standard deviations. It assumes without question that these are the same and that there is only uncertainty as to the means. In many problems however there will be *à priori* reasons for believing that there can be no significant difference in variability and it is in such cases that the test is applicable.

* See footnote to p. 197 above as to notation.

(*e*) Finally it is possible to make use of the P_A Tables in the following manner. Having observed the two samples, the single normal population of maximum likelihood is that associated with hypothesis (1) of the preceding paragraph, (*d*), namely in the present case a population with $a = 54{\cdot}543$ secs., $\sigma = 2{\cdot}123$ secs. It may now be asked whether the two series of observations are likely to have arisen independently in random sampling *if this had been the population sampled.* It is found that

$$\text{for I, } n = 30,\ M = -{\cdot}397,\ S = {\cdot}887,\ k = {\cdot}51,\ P_A = {\cdot}079,$$
$$\text{for II, } n = 40,\ M = -{\cdot}298,\ S = {\cdot}976,\ k = {\cdot}47,\ P_A = {\cdot}201.$$

That is to say in the most favourable light possible, the combined probability of occurrence of the two samples may be measured by the product ·079 × ·201 = ·016. For any other values assigned to a and σ the probability of the observed result measured in this manner will be less, and therefore we should feel considerable confidence in arguing that the hypothesis of a common origin for the samples is improbable. This method does take into account possible differences in variability, and used with discretion would seem to be of value. It would of course lead to some difficulty of interpretation should the resulting product of probabilities be, let us say, ·05 to ·20; if in the most favourable circumstances the hypothesis of a common origin is just possible, are we to discard it on the grounds that we have no right to postulate the most favourable conditions? A satisfactory answer to this question is probably impossible. At any rate the method is perhaps one which encourages an attitude of caution.

In the present example, however, whichever test be used, it is clear that we should conclude that process I does almost certainly lead to a really quicker timing than process II.

III. Sampling from a Rectangular Population.

(10) *Sampling Distributions of the Frequency Constants.*

The population, Π, may be defined by two quantities, the range w, and the distance of the centre of this range from a fixed origin, g. Then all values of the variable between $g - \frac{1}{2}w$ and $g + \frac{1}{2}w$ are of equal frequency of occurrence in Π. In a particular sample, Σ, consisting of n observations $X_1, X_2, \ldots X_n$, the range may be denoted by W and the distance of the mid-point or "centre" of this range from the fixed origin by G, that is to say the lowest and highest variates of the sample are $G - \frac{1}{2}W$ and $G + \frac{1}{2}W$ respectively. The fundamental space for samples of n drawn randomly from Π will be an n-dimensioned hypercube each of the sides of which is of length w, and having as a diagonal the line joining the points $(g - \frac{1}{2}w, \ldots g - \frac{1}{2}w)$ and $(g + \frac{1}{2}w, \ldots g + \frac{1}{2}w)$. As in the "normal" space, this diagonal forms the axis of the $n!$ similar regions of which one is that defined by the condition

$$x_1 \leqslant x_2 \leqslant x_3 \leqslant \ldots \leqslant x_n \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(ii } \textit{bis}\text{)},$$

where $x_s = X_s - g + \frac{1}{2}w$, the origin for the x's being taken at one corner of the cube.

The density field within the hypercube will be uniform, for the chance of drawing a sample with observations in the limits $X_1 \pm \frac{1}{2}dX_1, \ldots X_n \pm \frac{1}{2}dX_n$ is $dX_1, \ldots dX_n/w^n$. Making use of this density field, Hall* has shown that the distribution of the means of samples of n is given by a curve consisting of n connected arcs of degree $n-1$, having $(n-1)$-point contact at their joins. The distribution of standard deviations seems likely also to consist of a number of connected arcs.

In testing Hypothesis A we cannot make use of contours of equal density as these do not exist. Nor, except from the established custom of regarding them as the most important frequency constants, is there any reason why the mean and standard deviation of the sample should be chosen as the most suitable estimates of the population parameters. Consider the problem from the point of view of the criterion of likelihood. Suppose that there are reasons for believing that Σ has been drawn from *some* population following a rectangular distribution, and that an alternative to Π is a population Π' with centre at g' and range of w'. Then

$$\text{Likelihood of } \Pi/\text{Likelihood of } \Pi' = w'^n/w^n.$$

For a given Σ, the population of maximum likelihood is that for which w' is a minimum. But w' cannot be less than the sample range, W, and if it has this value clearly $g' = G$. It follows that in the fundamental hypercube, the contours upon which

$$\lambda = \text{Likelihood of } \Pi/\text{Likelihood of } \Pi' \text{ (max)}$$

is constant are the contours of constant range.

If therefore we make use of the criterion of likelihood in forming a judgment on Hypothesis A, we need consider only the range of the sample, and our confidence in the hypothesis will increase as $W \to w$. If $W > w$ or any of the observations in the sample fall outside the range of Π, then of course the hypothesis becomes impossible. Remembering that we are only considering cases where the sample has come from *some* rectangular population, the use of range alone as a criterion is seen to be perfectly reasonable. This limitation in the form of the sampled population appears more serious than that involved in the previous part of this paper†, simply because populations of normal or approximately normal form are found in experience to be common, so that we can frequently assume normality without great risk of error, while it is more difficult to conceive of problems in which a rectangular form of population can be assumed *à priori*. There is little doubt however that useful conclusions would be reached by applying the tests developed below to samples from any nearly symmetrical population with sharply limited range—that is to say, a population for which β_1 and β_2 lay in the neighbourhood of 0 and 1·8.

* *Biometrika*, Vol. XIX. pp. 240—244.

† There it was assumed that Σ came from *some* normally distributed or nearly normally distributed population.

We will now consider the distribution of W and G in sampling from II. In Figure 7, OI is of length w, and A and B represent the lowest and highest observations of the sample.

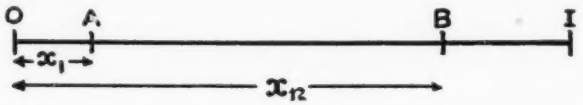


FIG. 7

The chance of obtaining one observation in an element dx centred at A, another in an element dx_n centred at B, and $n-2$ between these two values is

$$\frac{dx_1}{w} \times \left(\frac{x_n - x_1}{w}\right)^{n-2} \times \frac{dx_n}{w} = \left(\frac{x_n - x_1}{w^n}\right)^{n-2} dx_1\, dx_2.$$

Change the variables to $\begin{cases} G = \frac{1}{2}(x_1 + x_n) + g - \frac{1}{2}w, \\ W = x_n - x_1, \end{cases}$

so that $$dG\, dW = dx_1 dx_2.$$

Then the chance of obtaining a sample with centre and range in the limits $G \pm \frac{1}{2}dG$, $W \pm \frac{1}{2}dW$ is

$$\phi_1(G, W)\, dG\, dW \equiv k\, W^{n-2}/w^n \,.\, dG\, dW \ldots\ldots\ldots\ldots\ldots\text{(xxxvii)}.$$

Choosing k so that the integral of ϕ taken throughout the total range of variation is unity, it is found that $k = n(n-1)$. Hence

$$\phi_1(G, W) \equiv n(n-1)\, W^{n-2}/w^n \ldots\ldots\ldots\ldots\ldots\text{(xxxvii } bis\text{)}.$$

Integrating out for G between the limits $g - \frac{1}{2}(w - W)$ and $g + \frac{1}{2}(w - W)$ we obtain the chance, $\phi_2(W)\, dW$, of drawing a sample with range in the limits $W \pm \frac{1}{2}dW$ or

$$\phi_2(W) \equiv n(n-1)\frac{W^{n-2}}{w^{n-1}}\left(1 - \frac{W}{w}\right) \ldots\ldots\ldots\ldots\ldots\text{(xxxviii)}.$$

This is a Type I curve (and also part of an $(n-1)$th order parabola) with the mean at $(n-1)w/(n+1)$, and a range from $W = 0$ to w*. For the reasons given above we therefore take as our criterion the probability integral

$$P_W = \int_0^W \phi_2(W)\, dW = \{nw - (n-1)W\}\, W^{n-1}/w^n \ldots\ldots\ldots\text{(xxxix)}.$$

The chance of drawing a sample with its centre in the limits $G \pm \frac{1}{2}dG$ is obtained by integrating (xxxvii) between the limits for W of 0, and $w - 2\,|\,G - g\,|$† or

$$\phi_3(G) \equiv n\,(w - 2\,|\,G - g\,|)^{n-1}/w^n \ldots\ldots\ldots\ldots\ldots\text{(xl)},$$

a symmetrical curve made up of two portions of $(n-1)$th order parabolae placed "back to back" on the ordinate through $G = g$. The form of distribution is shown roughly in Figure 8.

* This expression for the special case of $n=2$ has been given by Borel: *Traité du Calcul des Probabilités*, Tome I, Fasc. 1, p. 20, 1925.

† Throughout this paper $|\,q\,|$ denotes that the expression q within the rules is to be given a positive sign.

The chance of drawing a sample with centre at a distance as great or greater than $|G-g|$ from the centre of the population will be

$$P_G = 2\int_G^{g+\frac{1}{2}w} n\{w-2(G-g)\}^{n-1}/w^n \,.\, dG = \{w-2|G-g|\}^n/w^n \quad \ldots(\text{xli}).$$

Finally we may find the distribution of $z' = (G-g)/\frac{1}{2}W$. Making the transformation $z' = 2(G-g)/W$, we have $\frac{1}{2}W\,dW\,dz' = dW\,dG$, and from (xxxvii),

$$\phi_4(z',\ W) \equiv \tfrac{1}{2}n(n-1)\ W^{n-1}/w^n \quad \ldots\ldots\ldots\ldots\ldots\ldots(\text{xlii}).$$

For a given value of z', W can increase from 0 until $\frac{1}{2}W + |G-g| = \frac{1}{2}w$, or $W = w/(1+|z'|)$. Integrating out for W between these limits, we have

$$\phi_5(z') \equiv \tfrac{1}{2}(n-1)\{1+|z'|\}^{-n} \quad \ldots\ldots\ldots\ldots\ldots\ldots(\text{xliii}).$$

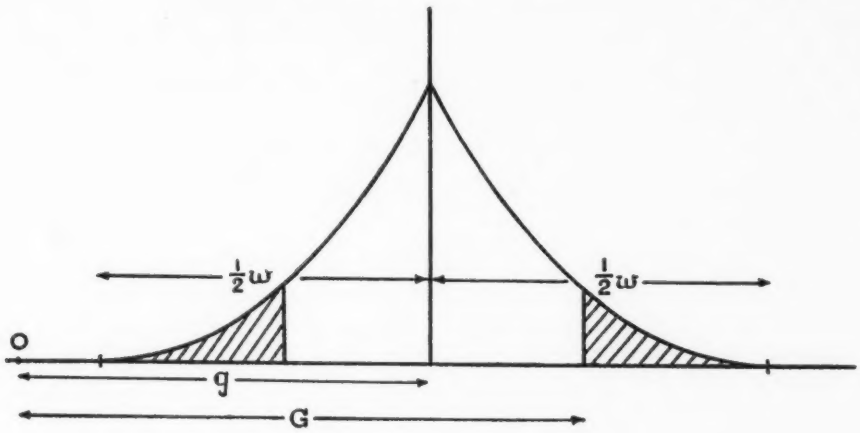


FIG. 8

This distribution of $z' = (G-g)/\frac{1}{2}W$ is of considerable interest. It corresponds to Student's distribution

$$f(z) = \text{constant} \times (1+z^2)^{-\frac{n}{2}}$$

for the ratio $z = m/s$ for samples from a normal population. z' is the ratio of the distance of the centre of the sample from the population centre to half the sample range, that is to say, like z it is the ratio of a measure of location to a measure of scale. The distribution of z' is independent of the population range w just as that of z is independent of the population standard deviation σ. It can be shown that z' is constant over parts of $n(n-1)$ primes passing through the centre of the fundamental hypercube and making an angle with its diagonal $x_1 = x_2 = \ldots = x_n$, which depends only on the value of z'. These primes correspond to the conical z contours in the normal space.

Suppose that we have a problem in which it is wished to know whether it is likely that a given sample $\Sigma\,(W, G)$ has been drawn from a rectangular population with centre at g, but for which the range, w, is unspecified. We are now testing Hypothesis B. Following the argument employed in dealing with normal populations, it is found, as shown in Figure 9, that the ratio of (1) Maximum

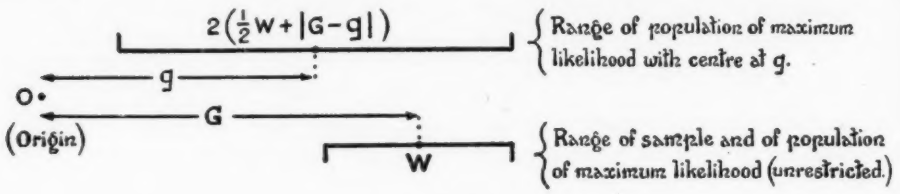


FIG. 9

likelihood for a rectangular population with centre at g, to (2) Maximum likelihood for any rectangular population, is

$$W^n/(W + 2\,|\,G - g\,|)^n = \{1 + |\,z'\,|\}^{-n}.$$

This expression is constant along the contours of z', and therefore leads to the probability integral of (xliii) as the criterion to employ in testing Hypothesis B, or

$$P_{z'} = (n-1)\int_{z'}^{\infty} (1+z')^{-n}\,dz' = (1+z')^{-n+1} \text{ .....(xliv).}$$

If the z' distribution is applicable only to exactly rectangular populations, its interest is mainly theoretical, but if it can be shown that it will represent adequately the distribution of the ratio (deviation in centre)/(half range) for small samples from a wider range of symmetrical populations, it will be of greater value. The test in this form or in its inverse form discussed below would enable an estimate to be made of the probable limits within which lies the mean of a symmetrical limited range population, from a knowledge of the two extreme individuals of the sample only. This it would do in the same way as Student's test does from a knowledge of the mean and standard deviation of the sample. Some further consideration of this point is being undertaken.

Fisher has given the limiting forms of the distributions of range (xxxviii) and centre (xl) when n is large*. He did not consider the case of small samples, but used his results as an illustration of the way in which the method of maximum likelihood picks out the "statistic" G as the estimate of the population mean which has the minimum standard error. If the estimate is to be based in this way on two observations only, it is of course necessary that the grouping should be fine, and that there should be no possibility that either of the extreme observations is subject to error of measurement or of record.

* *Phil. Trans.* A, Vol. 222, p. 348.

(11) *Solutions obtained by the Inverse Method.*

As a population is completely specified by the position of its centre g, and range w, the problem can be considered in the (g, w)-plane. This is represented in Figure 10.

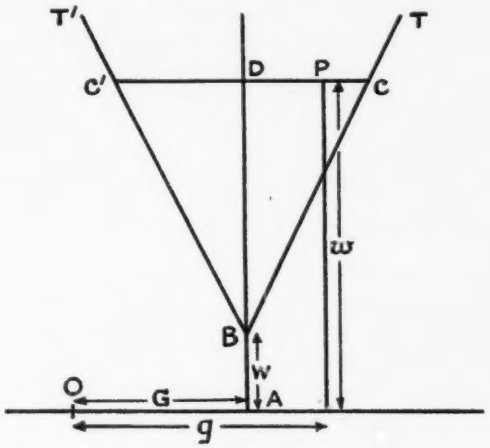


FIG. 10

B is the sample point (G, W), and P represents a population (g, w). For it to be possible that Σ has come from Π, it is necessary that

$$g - \tfrac{1}{2}w \leqslant G - \tfrac{1}{2}W \leqslant G + \tfrac{1}{2}W \leqslant g + \tfrac{1}{2}w,$$

or that

$$-\tfrac{1}{2}(w - W) \leqslant g - G \leqslant \tfrac{1}{2}(w - W),$$

or that

$$w \geqslant W + 2\,|\,g - G\,|.$$

Consequently the (g, w)-field is limited to the region inside the acute angle TBT', where $\angle TBD = \angle T'BD = \tan^{-1}\frac{1}{2}$. The likelihood of a rectangular population represented by P is $L \propto w^{-n}$, and we may give to the field a point-density $d = w^{-n}$. This is constant along lines such as CDC' parallel to the axis of g. L itself provides a criterion by which to judge the probability of Hypothesis A, but we obtain a more readily comprehensible scale by taking the ratio of (1) I', the integral of d throughout the portion of the field $TCC'T'$, to (2) I, the integral over the whole field TBT'. This gives, after simple integration,

$$P_w = I'/I = \{(n-1)\,w - (n-2)\,W\}\,W^{n-2}/w^{n-1} \quad \text{.....(xlv),}$$

and provides a measure not only of the probability that the range in the sampled population is $\geqslant w$, but also from the standpoint of likelihood it is the criterion to use as to the probability of Hypothesis A taken as a whole.

Finally we may consider the probability that the centre of the sampled population is such that $|\,g - G\,| \geqslant \alpha$, say. We shall obtain a measure of this by

integrating d over the two corresponding regions of the field, and dividing the result by I. It is found that

$$P_g = \left\{1 + \frac{2\,|\,g - G\,|}{W}\right\}^{-n+2}, \text{ where } |\,g - G\,| = \alpha \text{ .....(xlvi).}$$

Compare now (xlv) with (xxxix) and (xlvi) with (xliv), remembering that $z' = 2\,|\,G - g\,|/W$. It will be seen that by changing $n-1$ into n in the solutions of the inverse method, P_w becomes identical with P_W and P_g with $P_{z'}$. Consequently for a given sample (G, W) and hypothetical population (g, w), these pairs of criteria will give almost identical numerical values unless the sample is very small. The distributions of range and of the ratio z', (xxxviii) and (xliii), were in fact first obtained by one of us from the integrals P_w and P_g of the inverse method, making the change from $n-1$ into n from analogy with the solutions in the normal field!

(12) *Illustrations from Experimental Sampling.*

In practical application the weakness of the methods discussed—and indeed of any method of dealing with small samples—lies in the fact that while questioning the position (as measured by m or g) and the scale (by σ or w) of the sampled population, they provide no test of the shape of the distribution. They assume that the β_1 and β_2 of the population are known to be 0 and 3, or 0 and 1·8, as the case may be. In Section (8), analysing Church's samples of 10 from a skew population, we have considered from one direction what is the risk of error involved in making these assumptions. We shall now consider this from another direction.

Two sets of random samples from a rectangular population were obtained*:

(*a*) 500 samples of 4. Tippett's Random Sampling Numbers† were used, the distribution of the population being divided into 21 groups. As a check on the random nature of the sampling, the distribution of the total 2000 observations was compared with the "expected" distribution, namely one containing 95·24 observations in each of the 21 groups. Applying the (P, χ^2) test for goodness of fit, it was found that $\chi^2 = 16{\cdot}08$ and $P = {\cdot}712$.

(*b*) 540 samples of 10. These were obtained by drawing 10 coloured beads from a box containing 50 beads of each of 11 different varieties. The beads were replaced and shuffled after each draw of 10. The population was not therefore strictly "infinite," but as $M = 55n$ the effect is hardly likely to have been of great importance. Unfortunately it was not realised until afterwards that the sampling was biased because the various types of beads were sufficiently different in size to make the chances of drawing beads of the different colours not always equal to 1/11. The distribution of frequencies among the 11 groups for 4200 beads was 409, 446, 362, 404, 355, 407, 342, 422, 363, 375, 315. The "expected" number in each group is 381·8. Applying the (P, χ^2) test, it is found that

* (*a*) were drawn by Miss M. Page, (*b*) by Miss Page and two students of University College, London, Miss F. E. James and Mr T. Uchida. The work of computation has been mainly Miss Page's.

† Published as No. XV of the *Tracts for Computers*, Cambridge University Press.

$\chi^2 = 39{\cdot}71$ and $P = {\cdot}00002$. The samples must therefore be considered as drawn from some irregular and slightly skew population, and no very detailed analysis of the results seemed to be justified.

Taking the experimental sampling (a), we may first test the formulae of the preceding section.

The Distribution of Range, W.

In sampling, an observation fell into one or other of the 21 equal groups into which the population rectangle was divided. An exact range for the sample was not therefore obtainable, but it seemed that an adequate test could be obtained by taking the population range as $w = 21$, and by supposing that samples whose highest and lowest observations fell into:

(1) neighbouring groups, (2) groups 2 apart, (3) groups 3 apart, ... (20) groups 20 apart,

had ranges, W, falling into groups

(1) 0–1·5, (2) 1·5–2·5, (3) 2·5–3·5, ... (20) 19·5–21·0.

The theoretical distribution of W obtained from (xxxviii) is

$$y = 12\,(21\,W^2 - W^3)/(21)^4 \quad \text{.....(xlvii).}$$

Taking its integral over the corresponding grouping ranges for W, we have the set of observed and theoretical frequencies shown in Table IV. The (P, χ^2) test makes $\chi^2 = 13{\cdot}43$, which for $n' = 17$ gives $P = {\cdot}641$.

The Distribution of Centre, G.

The origin for G was taken at the centre of the middle or 11th population group; the values of G were either integral or contained half units.

A value of G, "0" was taken as lying in a group ranging from $-0{\cdot}25$ to $+0{\cdot}25$,

" " "0·5" " " " " $+0{\cdot}25$ " $+0{\cdot}75$,

" " "1·0" " " " " $+0{\cdot}75$ " $+1{\cdot}25$,

etc. .....

The theoretical distribution of G obtained from (xl) is

$$y = 4\,\{1 - 2\,|\,G\,|/21\}^3/21 \quad \text{.....(xlviii).}$$

Taking its integral for the required groups, we obtain the series of observed and theoretical frequencies for G given in Table IV. These results lead to $\chi^2 = 17{\cdot}31$, which for $n' = 21$ gives $P = {\cdot}633$.

For the mean and standard deviation of G, we have

	Observation	Theory	Standard Error
Mean	$-{\cdot}0950$	0	·1212
Standard deviation	2·6778	2·7111	·0857 *

* Approximately, on the assumption that the distribution of G is a normal curve.

TABLE IV. *Distribution of Frequency Constants in 500 Samples of 4 from a Rectangular Population.*

Frequencies for Range			Frequencies for Centre			Frequencies for ratio, z'			Frequencies for ratio, z		
W	Observed	Expected	G	Observed	Expected	z'	Observed	Expected	z	Observed	Expected ("normal" case)
			$-8{\cdot}75$ to $-9{\cdot}25$	1							
			$-8{\cdot}25$—	—							
			$-7{\cdot}75$—	—	6·7*						
			$-7{\cdot}25$—	3							
			$-6{\cdot}75$—	—							
			$-6{\cdot}25$—	5							
			$-5{\cdot}75$—	4	3·8	$-3{\cdot}00$ and less	6	3·9			
			$-5{\cdot}25$—	7	5·1	$-2{\cdot}75$ to $3{\cdot}00$	—				
			$-4{\cdot}75$—	6	6·9	$-2{\cdot}50$—	1	5·4			
0—$1{\cdot}5$	1	0·7	$-4{\cdot}25$—	10	8·9	$-2{\cdot}25$—	—		$-5{\cdot}0$ and less	3	1·6
$1{\cdot}5$—$2{\cdot}5$	1	2·3	$-3{\cdot}75$—	12	11·3	$-2{\cdot}00$—	2		$-4{\cdot}5$ to $5{\cdot}0$	1	
$2{\cdot}5$—	5	5·0	$-3{\cdot}25$—	9	14·1	$-1{\cdot}75$—	5	6·7	$-4{\cdot}0$—	2	
$3{\cdot}5$—	3	8·4	$-2{\cdot}75$—	18	17·4	$-1{\cdot}50$—	8		$-3{\cdot}5$—	1	1·9
$4{\cdot}5$—	10	12·3	$-2{\cdot}25$—	26	21·1	$-1{\cdot}25$—	4	6·0	$-3{\cdot}0$—	—	
$5{\cdot}5$—	25	16·7	$-1{\cdot}75$—	21	25·3	$-1{\cdot}00$—	16	9·3	$-2{\cdot}5$—	7	2·2
$6{\cdot}5$—	20	21·2	$-1{\cdot}25$—	27	30·0	$-0{\cdot}75$—	12	15·4	$-2{\cdot}0$—	3	4·4
$7{\cdot}5$—	25	25·7	$-0{\cdot}75$—	43	35·3	$-0{\cdot}50$—	25	27·4	$-1{\cdot}5$—	12	10·0
$8{\cdot}5$—	29	30·0	$-0{\cdot}25$—	40	41·2	$-0{\cdot}25$—	52	53·9	$-1{\cdot}0$—	20	25·3
$9{\cdot}5$—	40	33·9	$+0{\cdot}25$ to $-0{\cdot}25$	38	45·8	0 to $-0{\cdot}25$	121	122·0	$-0{\cdot}5$—	60	67·1
$10{\cdot}5$—	33	37·3	$+0{\cdot}75$ to $+0{\cdot}25$	38	41·2	0 to $+0{\cdot}25$	113	122·0	0 to $-0{\cdot}5$	139	137·5
$11{\cdot}5$—	45	39·9	$+1{\cdot}25$—	41	35·3	$+0{\cdot}25$—	65	53·9	0 to $+0{\cdot}5$	145	137·5
$12{\cdot}5$—	46	41·7	$+1{\cdot}75$—	33	30·0	$+0{\cdot}50$—	30	27·4	$+0{\cdot}5$—	72	67·1
$13{\cdot}5$—	47	42·3	$+2{\cdot}25$—	34	25·3	$+0{\cdot}75$—	14	15·4	$+1{\cdot}0$—	18	25·3
$14{\cdot}5$—	35	41·6	$+2{\cdot}75$—	15	21·1	$+1{\cdot}00$—	10	9·3	$+1{\cdot}5$—	7	10·0
$15{\cdot}5$—	32	39·4	$+3{\cdot}25$—	20	17·4	$+1{\cdot}25$—	5	6·0	$+2{\cdot}0$—	3	4·4
$16{\cdot}5$—	35	35·6	$+3{\cdot}75$—	15	14·1	$+1{\cdot}50$—	2	6·7	$+2{\cdot}5$—	3	2·2
$17{\cdot}5$—	28	29·9	$+4{\cdot}25$—	11	11·3	$+1{\cdot}75$—	1		$+3{\cdot}0$—	2	1·9
$18{\cdot}5$—$19{\cdot}5$	24	22·2	$+4{\cdot}75$—	4	8·9	$+2{\cdot}00$—	3		$+3{\cdot}5$—	1	
$19{\cdot}5$—$21{\cdot}0$	16	13·9	$+5{\cdot}25$—	8	6·9	$+2{\cdot}25$—	2	5·4	$+4{\cdot}0$—	—	
			$+5{\cdot}75$—	3	5·1	$+2{\cdot}50$—	—		$+4{\cdot}5$ to $5{\cdot}0$	1	1·6
			$+6{\cdot}25$—	4	3·8	$+2{\cdot}75$ to $3{\cdot}00$	—		$+5{\cdot}0$ and more	—	
			$+6{\cdot}75$—	1	6·7*	$+3{\cdot}00$ and more	3	3·9			
			$+7{\cdot}25$—	2							
			$+7{\cdot}75$ to $+7{\cdot}25$	1							
Total	500	500·0	Total	500	500·0	Total	500	500·0	Total	500	500·0

* Theoretical frequencies from $G = 6{\cdot}25$ to $10{\cdot}5$, or tail of curve.

The standard deviation of the sampled population is $21/(2\sqrt{3}) = 6{\cdot}0622$; consequently the standard deviation of the mean in samples of 4 is $\sigma/\sqrt{n} = 3{\cdot}0311$, illustrating Fisher's point referred to above that G provides a better estimate of the population centre or mean than $\bar{X}$, the sample mean.

The Distribution of $z' = 2(G-g)/W$.

Using (xliii), we see that the distribution of z' is

$$y = \tfrac{3}{2}(1+|z'|)^{-4} \quad \text{.....(xlix).}$$

Integrating this expression within the grouping limits shown, we obtain the observed and theoretical frequencies given in Table IV. Here $\chi^2 = 12{\cdot}64$ and $n' = 14$, so that $P = {\cdot}477$.

The agreement of observation with theory for these three distributions is therefore very satisfactory. We may consider now the further problem. There exist a great variety of limited range symmetrical (and unimodal) distributions with values of β_2 lying between 3·0 and 1·8; it is very unlikely that, however inadequate our information, we should suppose that the population from which a sample has been drawn to be normal when in fact it was rectangular. But we can obtain some idea of the degree to which the normal criteria will fail for samples from populations with intermediate forms—say with β_2 2·0 to 2·6—by testing the adequacy of the normal theoretical distributions in this extreme case of $\beta_2 = 1{\cdot}8$.

Take first the distribution of range. For a rectangular distribution $w = 2\sqrt{3}\sigma$; it follows from (xxxviii) that the frequency constants of the sample range, W, are:

$$\left.\begin{aligned} &\text{Mean} = \frac{n-1}{n+1}w = 2\sqrt{3}\frac{n-1}{n+1}\sigma \\ &\text{Standard Deviation} = \frac{w}{n+1}\sqrt{\frac{2(n-1)}{n+2}} = \frac{2\sigma}{n+1}\sqrt{\frac{6(n-1)}{n+2}} \\ &\beta_1 = \frac{2(n+2)(n-3)^2}{(n-1)(n+3)^2}; \quad \beta_2 = \frac{6(n+2)(n^2-2n+3)}{(n-1)(n+3)(n+4)} \end{aligned}\right\}.$$

The ratios of mean range and of standard deviation of range to the standard deviation of the population, and the β_1 and β_2 of the range distribution for samples from a normal population, are known*. These values are compared with the true values in Table V, for $n = 3$, 4, 6, 10 and 20. For the smaller samples there is a fair degree of correspondence in the mean range, but as n increases the ratio for the rectangular distribution tends to $2\sqrt{3} = 3{\cdot}4641$ while that for the normal distribution increases slowly but without limit. Even at $n = 3$, the standard deviation ratios have little correspondence, while the β_1 and β_2 values show no agreement, except in giving increasing skewness as n increases.

* L. H. C. Tippett, *Biometrika*, Vol. XVII. pp. 364—387. The values given in the table were actually taken from the summary given in *Biometrika*, Vol. XVIII. p. 192.

* Theoretical frequencies from $G = 6{\cdot}25$ to $10{\cdot}5$, or tail of curve.

TABLE V.

Comparison of Frequency Constants of Range.

	Population	Size of Sample				
		3	*4*	*6*	*10*	*20*
Mean W/σ ...	Rectangular	1·732	2·078	2·474	2·834	3·134
	Normal	1·693	2·059	2·534	3·078	3·735
Standard Deviation of W/σ	Rectangular	·775	·693	·553	·386	·217
	Normal	·888	·880	·848	·797	·729
β_1 ...	Rectangular	0·000	·082	·356	·773	1·265
	Normal	·417	·273	·189	·156	·161
β_2 ...	Rectangular	2·143	2·357	2·880	3·648	4·569
	Normal	3·286	3·188	3·170	3·22	3·26

Consider next what would be the result of supposing the sampled population to have been normal, and applying Student's z-test to measure the probability of Hypothesis B. Does the distribution of $z = m/s$ for these samples follow approximately the law $f(z) = \text{constant} \times (1 + z^2)^{-\frac{n}{2}}$? A comparison of observation and normal theory is given in the fourth section of Table IV. It is found that $\chi^2 = 10{\cdot}77$, which for $n' = 10$ gives $P = {\cdot}293$. The main discrepancy, which is not serious, lies in some large negative values of the observed z's. The observed standard deviation is $\sigma_z = 1{\cdot}3030$ against the normal value of $1/\sqrt{n-3} = 1{\cdot}0000$. As the standard error of σ_z in the normal case is infinite, owing to the 4th moment of the z-distribution being infinite, it is not easy to estimate the significance of this difference. Judging only from the χ^2-test, however, Student's distribution appears quite adequate to describe the variation in z, as far as the evidence from these 500 samples goes.

It must be remembered, however, that even if z calculated for samples from a rectangular distribution follows Student's law, P_z has changed its meaning as a criterion; it no longer enables us to reject or accept Hypothesis B upon the same logical basis. As we have seen, a more suitable ratio to employ would now be $z' = 2(G - g)/W$. But is z highly correlated with z'? If this were so, in judging from z we should reach results not very different from those based on z'. We have calculated the coefficient of correlation between z and z' for the 500 samples; it has the very high value of $+ {\cdot}9756$. These results therefore suggest that for samples of 4 from a rectangular population, (*a*) the distribution of $z = m/s$ follows Student's law with sufficient approximation for ordinary purposes, and (*b*) the criterion P_z will be closely equivalent to $P_{z'}$. Consequently, if we assume

erroneously that the sampled population is normal when in fact it has a β_1 and β_2 on or close to the stretch of the β_2 axis lying between $\beta_2 = 3{\cdot}0$ and $1{\cdot}8$, we shall not be led into serious errors of judgment by using Student's test. The results for samples of 10 considered below suggest that for this case also point (a) holds good, and since as n increases the distribution of m/s tends to that of m/σ or to a normal curve with standard deviation $\sigma/\sqrt{n}$ whatever be the form of the population sampled, the most crucial test of correspondence (except for the cases $n = 2$ and 3) would appear to have been made at $n = 4$. On the other hand, the correlation between z and z' which is perfect at $n = 2$ will decrease as n increases, and the mean and standard deviation of the sample are less and less influenced by the values of the two extreme observations. The interpretation of the z-test will therefore become more and more doubtful.

TABLE VI.

Distribution of z in 540 samples of 10.

Central values of z	Less than $-1{\cdot}80$	$-1{\cdot}75$	$-1{\cdot}65$	$-1{\cdot}55$	$-1{\cdot}45$	$-1{\cdot}35$	$-1{\cdot}25$	$-1{\cdot}15$	$-1{\cdot}05$	$-{\cdot}95$	$-{\cdot}85$	$-{\cdot}75$
Observation ...	—	1	—	—	1	1	1	1	3	1	5	5
Normal Theory	0·1	0·1	0·1	0·1	0·2	0·4	0·6	0·9	1·5	2·5	4·2	6·8
z	$-{\cdot}65$	$-{\cdot}55$	$-{\cdot}45$	$-{\cdot}35$	$-{\cdot}25$	$-{\cdot}15$	$-{\cdot}05$	$+{\cdot}05$	$+{\cdot}15$	$+{\cdot}25$	$+{\cdot}35$	$+{\cdot}45$
Observation ...	11	13	25	29	44	53	71	55	86	47	29	18
Normal Theory	10·9	16·8	25·1	35·4	46·3	56·1	61·9	61·9	56·1	46·3	35·4	25·1
z	$+{\cdot}55$	$+{\cdot}65$	$+{\cdot}75$	$+{\cdot}85$	$+{\cdot}95$	$+1{\cdot}05$	$+1{\cdot}15$	$+1{\cdot}25$	$+1{\cdot}35$	Greater than $+1{\cdot}40$		Total
Observation ...	19	11	1	2	2	3	1	—	1	—		540
Normal Theory	16·8	10·9	6·8	4·2	2·5	1·5	0·9	0·6	0·4	0·6		540·0

We shall now consider briefly the experimental sampling (b), with 540 samples of 10. As stated above, no very detailed analysis is justifiable. Since the population was only divided into 11 groups, it was not possible to obtain satisfactory distributions of W, G or z'. The mean, standard deviation and ratio z were, however, calculated. Owing to the biased sampling, the population could not be considered as exactly rectangular. In order to correct to some extent the error involved, the means, m, were measured not from the centre of the middle or 6th population group, but from a point $-{\cdot}19$ from this, corresponding more nearly to the mean as observed from the samples.

The observed distribution of z and the corresponding theoretical frequencies are shown in Table VI. The fit is bad, giving $\chi^2 = 27{\cdot}16$, which for $n' = 16$ makes $P = {\cdot}027$. A comparison of the frequencies shows that this is due to a tremendous

contribution of 15·94 to χ^2 for the group $z = +·1$ to $+·2$. But beyond this, the observed distribution of z is somewhat skew, having $\beta_1 = ·395$ and $\beta_2 = 4·988$. Now this cannot be due to the error involved in assuming that the z of samples from a rectangular population follows the distribution appropriate for a normal population, for in both cases the distribution of z must be symmetrical. It is probable that the asymmetry as well as the great discrepancy in the group $z = +·1$ to $+·2$ results from the irregularity of the bead sampling combined with some effect on the grouping of z introduced by the correction to m referred to above.

Using the only other measure of comparison, the standard deviation of z, we find:

from observations $\sigma_z = ·3707$; from theory $\sigma_z = ·3780$,

with a standard error which would, if the distribution of σ_z were normal, be ·0113. It appears therefore that as far as the scale goes, the observed z distribution is represented by the "normal" value of σ_z, or $1/\sqrt{n-3}$. Further sampling would be necessary to determine the form of the curve*, but as the probability integrals of symmetrical curves with the same standard deviation change very slowly with β_2, it seems very probable that for $n = 10$ Student's distribution will still represent adequately the distribution of z from a rectangular population. The interpretation of P_z in testing Hypothesis B is again of course doubtful.

While the distributions of z appear in both these series of samples to be represented approximately by Student's curve, the normal P_λ contours are found to be quite inadequate. This is partly because the regression of s on m is no longer linear, and partly because the distribution of the s-margin cannot be represented even approximately by the normal s-distribution.

The standard deviation of the variance, s^2, is given by

$$\sigma_{s^2} = \sigma^2 \frac{(n-1)}{n^{\frac{3}{2}}} \sqrt{\beta_2 - 3 + \frac{2n}{n-1}},$$

where β_2 refers to the sampled population. Further $\sigma_s = \sigma_{s^2}/2\sigma$ fairly closely for large samples, though with increasing approximation as n decreases. It follows

TABLE VII.

	Observations from Rectangular Population					Theory for Normal Population				
n	Mean s/σ	σ_s/σ	β_1	β_2	*Sk.*	Mean s/σ	σ_s/σ	β_1	β_2	*Sk.*
4	·822	·275	·006	2·379	− ·074	·798	·337	·236	3·108	+ ·270
10	·932	·148	·070	2·761	− ·174	·923	·220	·063	3·011	+ ·128

that σ_s is very sensitive to changes in β_2, and the drop from 3·0 to 1·8 in β_2 completely alters its value. Also, as the results in Table VII show, the observed

* In view of the complex nature of the distribution of m as found by Irwin and Hall (*Biometrika*, Vol. XIX. pp. 225 and 240) it seems unlikely that the theoretical distribution of z would be easy to find.

distributions of s for the samples from the rectangular population are skew in the opposite sense to those from the normal population; the skewness increases with n.

In view of this complete lack of correspondence, except perhaps for the mean s, it is not surprising to find that the P_λ contours are here of little value. Placing down upon each (m, s)-distribution three contours, the results of Table VIII below were obtained. The observed frequencies are in all cases less than those expected on normal theory, so that in using the P_λ contours we should be inclined to underestimate the significance of a divergence in the sample point. At $\beta_2 = 1{\cdot}8$ we have reached a position quite out of touch with the normal field; it is hoped shortly to provide information on the position at an intermediate point, say $\beta_2 = 2{\cdot}5$.

TABLE VIII.

$n=4$			$n=10$		
k	Observed frequency outside contour	$P_\lambda \times 500$	k	Observed frequency outside contour	$P_\lambda \times 540$
$\cdot 70$	138	196·5	$\cdot 50$	179	271·9
$1{\cdot}00$	48	68·8	$\cdot 70$	19	33·8
$1{\cdot}50$	11	12·1	$\cdot 90$	2	4·2

IV. Sampling from an Exponential Population.

(13) *Sampling Distributions of Frequency Constants.*

The frequency distribution of the population may be written

$$y = \frac{1}{\sigma} e^{-\frac{X-b}{\sigma}} \quad \text{.....(l),}$$

where b is the distance from the origin to the start of the curve, $b + \sigma$ the distance from the origin to the mean, and σ the standard deviation of the curve. In the fundamental space of Π, we take rectangular axes through the point $(b, b, \ldots b)$, and writing $x = X - b$, the field in which Σ may lie is that for which all the x's are positive. The appropriate density with which to fill this field will be

$$D = \sigma^{-n} e^{-\frac{S(x)}{\sigma}} = \sigma^{-n} e^{-\frac{n\bar{x}}{\sigma}} \quad \text{.....(li),}$$

where $\bar{x}$ is the distance of the sample mean from the start of the population curve. The contours of equal density will be the m contours, which are primes in the n-fold space lying at right angles to the diagonal axis

$$x_1 = x_2 = \ldots = x_n \quad \text{.....(iii } bis\text{).}$$

We shall now find the contours in this space upon which the ratio

$$\lambda = \text{Likelihood of } \Pi / \text{Likelihood of } \Pi' \text{(max)}$$

is constant. For an exponential population, Π', with start at b' and standard deviation σ', we have

$$L \propto \sigma'^{-n} e^{-\frac{n(\bar{X}-b')}{\sigma'}} \quad \text{.....(lii).}$$

Whence $\frac{\partial L}{\partial b'} \propto n\sigma'^{-n-1} e^{-\frac{n(\bar{X}-b')}{\sigma'}}$, showing that for a given value of σ', L is a maximum when $\bar{X}-b'$ is as small as possible. Clearly this will occur when $b'=X_1$, the lowest of the sample variates, for it is impossible for Σ to have come from Π' if $X_1 < b'$.

Further $\frac{\partial L}{\partial \sigma'} \propto n\sigma'^{-n-2} e^{-\frac{n(\bar{X}-b')}{\sigma'}} \{\bar{X}-b-\sigma'\}$. This expression vanishes when

(*a*) $\sigma'=0$ making $L=0$.
(*b*) $\sigma'=\infty$ „ $L=0$.
(*c*) $\sigma'=\bar{X}-b'$ „ $L \propto (\bar{X}-b')^{-n} e^{-n}$.

The last is the maximum solution; it follows that for a given sample, the exponential population of maximum likelihood is that for which

$$b'=X_1=b+x_1; \quad \sigma'=\bar{X}-X_1=\bar{x}-x_1.$$

Hence the ratio λ is given by

$$\lambda=\left(\frac{\bar{x}-x_1}{\sigma}\right)^n e^{-\frac{n(\bar{x}-\sigma)}{\sigma}} \quad \text{.....(liii).}$$

This result shows that the method of likelihood picks out $\bar{x}$ the mean, and x_1 the lowest observation, as the two descriptive measures of the sample; that is to say it suggests that we should take $X_1=b+x_1$ and $l=\bar{X}-X_1=\bar{x}-x_1$ as the sample estimates of the population parameters b and σ respectively (see Figure 12). In the fundamental space, λ is constant upon the contours

$$(x_1+l)/\sigma-\log_e(l/\sigma)=k=1-(1/n)\log_e\lambda \quad \text{.....(liv),}$$

and this will be the system to use in testing Hypothesis A according to the criterion of likelihood. The problem can be referred to the (x_1, l)-plane, and it is necessary to find the appropriate density, d, for this field. This can be obtained most readily by obtaining first the $(x_1, \bar{x})$-field.

The chance of drawing a sample from Π with the n observations in the limits $x_1 \pm \frac{1}{2}dx_1, \ldots x_n \pm \frac{1}{2}dx_n$ is

$$C dx_1 \ldots dx_n = \sigma^{-n} e^{-\frac{n\bar{x}}{\sigma}} dx_1 \ldots dx_n.$$

Take now as variables

$$\begin{cases} z_0 = x_1 + x_2 + \ldots + x_n, \\ z_1 = x_1, \text{ the lowest observation of sample,} \\ z_3 = x_3, \ z_4 = x_4, \ldots z_n = x_n. \end{cases}$$

Hence

$$C dx_1 dx_2 \ldots dx_n = C dz_0 dz_1 dz_3 \ldots dz_n = \sigma^{-n} e^{-\frac{z_0}{\sigma}} dz_0 dz_1 dz_3 \ldots dz_n.$$

For a given set of values of $z_0, z_1, z_4, \ldots z_n$, the variable z_3 cannot be less than z_1 and takes its highest value when $x_2 = x_1 = z_1$, and consequently

$$2x_1 + x_3 + x_4 + \ldots + x_n = n\bar{x}, \text{ or } z_3 = z_0 - 2z_1 - z_4 \ldots - z_n = u_1.$$

Hence

$$C dz_0 dz_1 dz_4 \ldots dz_n = dz_0 dz_1 dz_4 \ldots dz_n \sigma^{-n} e^{-\frac{z_0}{\sigma}} \int_{z_1}^{u_1} dz_3$$

$$= dz_0 \ldots dz_n \sigma^{-n} e^{-\frac{z_0}{\sigma}} (z_0 - 3z_1 - z_4 \ldots - z_n).$$

Now for given values of $z_0, z_1, z_3, \ldots z_n$, the variable z_4 can vary from z_1 to $u_2 = z_0 - 3z_1 - z_3 \ldots - z_n$, corresponding to the case where both x_2 and x_3 have their lowest possible values of x_1.

Hence

$$C dz_0 dz_1 dz_3 \ldots dz_n = dz_0 \ldots dz_n \sigma^{-n} e^{-\frac{z_0}{\sigma}} \int_{x_1}^{u_2} (z_0 - 3z_1 - z_4 \ldots - z_n)\, dz_4$$

$$= dz_0 \ldots dz_n \sigma^{-n} e^{-\frac{z_0}{\sigma}} \tfrac{1}{2} (z_0 - 4z_1 - z_3 \ldots - z_n)^2.$$

Proceeding in this way we can integrate out successively for all the variables $z_3, z_6, \ldots z_n$, and shall be left with

$$C dz_0 dz_1 = dz_0 dz_1 \sigma^{-n} e^{-\frac{z_0}{\sigma}} (z_0 - nz_1)^{n-2}/(n-2)!.$$

This process of integration will have covered only $1/n$ of the complete fundamental space, namely that portion in which $x_1 = z_1$ is the lowest observation. There will be $n-1$ similar regions in which $x_2, x_3, \ldots$ or x_n is the lowest. To obtain the chance that the mean of the sample lies in the limits $\bar{x} \pm \frac{1}{2} d\bar{x}$, and that the lowest value of n undifferentiated observations lies in the limits $x_1 \pm \frac{1}{2} dx_1$, we must multiply the above expression by n and write $z_0 = n\bar{x}$, $z_1 = x_1$, giving

$$\phi_1(\bar{x}, x_1)\, d\bar{x}\, dx_1 \equiv \sigma^{-n} (n^n/(n-2)!)\, (\bar{x} - x_1)^{n-2} e^{-\frac{n\bar{x}}{\sigma}} d\bar{x}\, dx_1 \quad \ldots\ldots\ldots(\text{lv}).$$

From (lv), integrating for x_1 between the limits 0 and $\bar{x}$, we find the distribution of means, $\bar{x}$, or

$$\phi_2(\bar{x}) \equiv \sigma^{-n} (n^n/(n-1)!)\, \bar{x}^{n-1} e^{-\frac{n\bar{x}}{\sigma}} \quad \ldots\ldots\ldots\ldots\ldots\ldots(\text{lvi}),$$

a Type III curve corresponding to a special case of the result obtained by Church*.

Next, upon integrating out for $\bar{x}$ between the limits x_1 and ∞, we obtain for the distribution of the lowest observation in the sample,

$$\phi_3(x_1) \equiv (n/\sigma) e^{-\frac{nx_1}{\sigma}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(\text{lvii}),$$

which is another exponential curve, with standard deviation σ/n.

* *Biometrika*, Vol. XVIII. p. 336, giving the distribution of means in samples of n from a Type III population.

Putting in (lv), $\bar{x} = l + x_1$, we obtain for the frequency function of x_1 and l,

$$\phi_4(x_1, l) \equiv \sigma^{-n}(n^n/(n-2)!)\, l^{n-2} e^{-\frac{n(l+x_1)}{\sigma}} \quad \text{.....(lviii).}$$

If then we know that the sample Σ has been drawn from *some* exponential population and we wish to judge the probability that it has been taken from Π, the suitable criterion to take is P_λ, or the integral of a density $d = \phi_4(x_1, l)$ taken over the portion of the (x_1, l)-plane lying outside the member of the system (liv) passing through the sample point, (x_1, l).

The (x_1, l)-plane is represented in Figure 11, OA, OB being the axes of l and x_1. G is the point $x_1 = 0$, $l = \sigma$, representing the sample for which Π would be the population of maximum likelihood. By differentiating (liv) it is seen that

$$\frac{dx_1}{dl} = \frac{\sigma - l}{l},$$

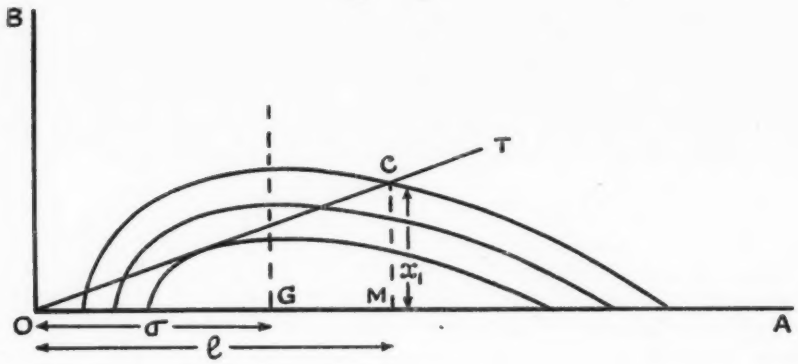


FIG. 11

and consequently all contours cut a line $l =$ constant at the same angle. The differential coefficient is zero along the line $l = \sigma$, and for very divergent contours tends to $+\infty$ and -1 at the two points of intersection with the axis of l. C represents a sample Σ; P_λ is the integral of the field taken outside the curve through C. We have not attempted to compute this.

Consider now Hypothesis B. Having a sample Σ, with observations $X_1, X_2, \ldots X_n$, which we have reason to believe has been drawn from *some* exponential population, we wish to measure the probability that it has been drawn from a population whose "start" or lower limit is at a distance b from the origin. The parameter, σ, of the population is unspecified. Then, as before, we consider the ratio of

(1) the maximum likelihood for an exponential population starting at b,

to (2) the maximum likelihood for any exponential population.

The former corresponds to the population for which $b' = b$, $\sigma' = \bar{X} - b$, the latter for which $b' = X_1$, $\sigma' = \bar{X} - X_1$, as found from (lii).

The position is shown diagrammatically in Figure 12, where the observations of the sample are shown as black circles, and the exponential curves of the two hypothetical populations have been dotted in. The ratio of the two likelihoods becomes

$$(\bar{X}-b)^{-n}e^{-n}/(\bar{X}-X_1)^{-n}e^{-n}=\left(1+\frac{x_1}{l}\right)^{-n}.$$

This is constant along the radiating lines, such as OCT in Figure 11, upon which the ratio $z''=x_1/l=\dfrac{X_1-b}{l}$ is fixed. We may therefore take these as the appropriate contours to use in testing Hypothesis B. It remains to find the distribution of z''. Writing in (lviii) $x_1=lz''$, we have

$$\phi_5(z'', l)\,dz''\,dl \equiv \sigma^{-n}(n^n/(n-2)!)\,l^{n-1}e^{-\frac{nl(1+z'')}{\sigma}}\,dz''\,dl \quad \text{.....(lix).}$$

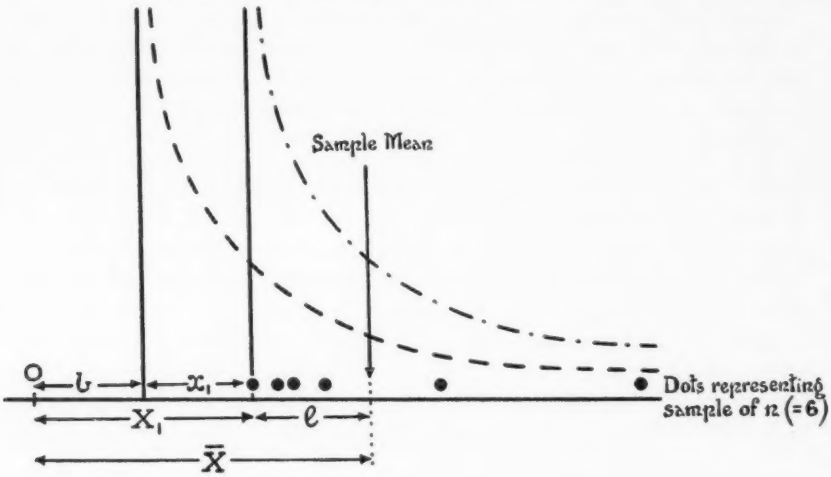


FIG. 12

This must be integrated for values of l between 0 and ∞. Writing

$$t=nl(1+z'')/\sigma,$$

we find

$$\phi_6(z'')\,dz''=(1+z'')^{-n}dz''\int_0^\infty t^{n-1}e^{-t}dt/(n-2)!,$$

or

$$\phi_6(z'')\equiv(n-1)(1+z'')^{-n} \quad \text{.....(lx).}$$

It is easily shown that

$$\text{Mean } z''=1/(n-2).$$

$$\text{Standard Deviation of } z''=\sqrt{(n-1)/(n-2)^2(n-3)}.$$

Further we obtain

$$P_{z''}=\int_{z''}^\infty (n-1)(1+z'')^{-n}\,dz''=(1+z'')^{-n+1} \quad \text{.....(lxi),}$$

which may be used as a criterion for testing Hypothesis B. It is comparable with and very similar to P_z, (xi), and $P_{z'}$, (xliv), corresponding to the normal and rectangular cases. z, z' and z'' are all ratios of a deviation in the measure of position to the measure of scale obtained from the sample.

These results which have been expressed in terms of x_1 and l are again what may be termed the "ideal" results for random sampling. Unfortunately in practice two difficulties often arise:

(*a*) There may be slight errors in observation, or variation in the character measured in an individual, which are not due to chance fluctuations.

(*b*) The unit of measurement or the grouping unit with which data have been tabled may be too coarse for the value of x_1 to be determined with adequate precision.

Under such conditions a method which provides a criterion depending upon the measure of a single variate, such as x_1, will be at a disadvantage compared with one which depends upon an average value, such as the mean or the standard deviation. In the latter case the effect of a few badly measured observations, a few disturbed variations or of coarse grouping is likely to cancel out. There are, however, we believe, problems to which the methods of the present section can be applied, and one of these in connection with random intervals we shall take as an illustration in the section that follows.

It has not seemed necessary to develop the inverse solution considered in dealing with the normal and rectangular populations, but the working would be quite straightforward.

(14) *Test for Random Intervals.*

Whitworth, in *Choice and Chance**, has shown that if the intervals of time between events which happen once on the average in time c may be considered as random, then the chance after an event has happened that another will not occur in an interval t is given by $e^{-\frac{t}{c}}$. If $y=f(t)$ represents the frequency distribution of the random intervals between the events, then

$$e^{-\frac{t}{c}} = \int_t^\infty f(t)\,dt, \text{ or } f(t) = \frac{1}{c}e^{-\frac{t}{c}}.$$

Morant† has extended this idea to the case where there is necessarily a closed interval of length b after an event before another can occur. The frequency distribution of intervals will now be

$$y = \frac{1}{c}e^{-\frac{t-b}{c}} \quad \text{.....(lxii).}$$

Suppose that we have $v=n+1$ consecutive observations of the occurrence of an event, recorded at times $t_1, t_2, t_3, \ldots t_v$, and that the n intervals or differences between the t's are $i_1, i_2, \ldots i_n$. Let i_s be the shortest of these intervals. Let us

* See p. 200, Prop. LI.

† *Biometrika*, Vol. XIII. pp. 309—337.

imagine in the first place that we have reason to believe that the intervals are random, but that there is between each a closed period of length b. We wish to test from the observations whether this hypothesis as to b is reasonable, but are not interested in the average time c. The problem is exactly that of testing Hypothesis B. With the previous notation

$$\begin{cases} x_1 &= i_s - b, \\ l &= (i_1 + i_2 + \ldots + i_n)/n - i_s = (t_v - t_1)/n - i_s, \\ z'' &= x_1/l, \\ P_{z''} &= (1 + z'')^{-n+1}. \end{cases}$$

In following this method it is important to be clear exactly what interpretation may be placed on the value found for $P_{z''}$. A large value (say $1 > P_{z''} > 0{\cdot}5$) will not imply that the intervals (less b) are random. This result will follow whatever be the form of distribution of intervals, provided only that $i_s - b$ is small compared to the average interval. A large $P_{z''}$ merely suggests that *if* the intervals (less the closed period) are random, then it is reasonable to suppose that the closed period is b. On the other hand, a very small value of $P_{z''}$ suggests that either the supposed closed period b is incorrect, or that the intervals (less b) are not random. In cases where it is clear *à priori* that there can be no closed period, then a small $P_{z''}$ certainly suggests that the intervals are not random.

If however it is possible to observe a number, say N, of series of n consecutive intervals, we may calculate z'' in each case and compare the distribution of these N values of z'' with the theoretical frequency law (lx). If there is agreement, then not only is it probable that the hypothesis as to b is correct, but also that the intervals (less b) are random. As z'' is independent of c, the average "rate of happenings" of the events, it follows that if b is unchanged—and in particular is zero—we may combine together in the test, the z'''s of different series of n observations, even though this rate may change from series to series. The test is also applicable, however small n may be. The practical importance of this will be illustrated below.

Through the kindness of Miss E. M. Newbold and Professor Greenwood, we have been able to make use of some of the original data regarding Factory Accidents used by the former in Report No. 34 of the Industrial Fatigue Research Board*. We may ask what answer the z''-test gives to the following question: Are the lengths of interval between consecutive accidents incurred by the same individual worker distributed randomly? For this purpose use has been made of the cards of male workers in a Chocolate Factory during the year 1923†. The time of accident was recorded to the day and even hour, and by making allowance

* "A Contribution to the Study of the Human Factor in the Causation of Accidents," 1925. For a later paper by the same writer, see *Journal of Roy. Stat. Soc.* 1927, p. 487.

† These are the cards of the series M 1 of page 11 of the I.F.R.B. Report, No. 34. The injuries were mainly of minor character, i.e. "cut finger," "abrasion," "splinter," etc. attended to in the Ambulance Room of the Factory, and did not lead to absence from work.

for Sundays, Bank Holidays, and as far as possible extra days of absence, sickness, etc.*, it was possible to obtain fairly accurately the number of days which each worker was exposed to risk between accidents. As Miss Newbold has shown, the liability to accident varies considerably among the workers, varying from none to fifteen accidents in the year for the cases examined, but the ratios z'', which are independent of this liability or rate, can be combined together.

The cards have been analysed for 50 workers; of these, 25 incurred only three accidents in the year, giving two intervals or the minimum to which the test can be applied.

TABLE IX.

Intervals between Accidents. Distributions of z''.

$n=2$			$n=4$			$n=6$		
z''	Observation	Theory	z''	Observation	Theory	z''	Observation	Theory
0— ·5	20	16·7	0— ·1	4	6·2	0—·1	4	7·6
·5— 1·0	7	8·3	·1— ·2	8	4·3	·1—·2	10	4·4
1·0— 1·5	4	5·0	·2— ·3	4	3·1	·2—·3	5	2·6
1·5— 2·0	2	3·3	·3— ·4	1	2·3	·3—·4	—	1·7
2·0— 2·5	2	2·4	·4— ·5	1	1·7	·4— ·5	—	1·1
2·5— 3·0	2·5	1·8	·5—1·0	3	4·3	Greater than ·5	1	2·6
3·0— 4·0	5	2·5	1·0—1·5	3	1·5			
4·0— 5·0	1·5	1·7	1·5—2·0	1	0·7	—	—	—
5·0— 6·0	—	1·2	2·0—5·0	—	0·8	—	—	—
6·0— 7·0	1	0·9	5·0—8·0	—	0·1	—	—	—
7·0— 8·0	—	0·7	—	—	—	—	—	—
8·0— 9·0	—	0·6	—	—	—	—	—	—
9·0—10·0	—	0·4	—	—	—	—	—	—
Greater than 10·0	5	4·5	—	—	—	—	—	—
			—	—	—	—	—	—
Total	50	50·0	Total	25	25·0	Total	20	20·0
Mean z''	—	—		·460	·500		·179	·250
Difference in terms of standard error				−·23			−·99	

In the above table are given the calculated values of z'' for (a) the 50 cases, taking the first three accidents or $n=2$, (b) the 25 of these for which there were at least five accidents ($n=4$), and (c) the 20 of these for which there were at least seven accidents ($n=6$). The intervals chosen were the first two, first four and first six of the year. It is clear that as the records were restricted to a year, they do not contain a completely random selection of intervals, at any rate for the workers

* As the dates of absence were not specified exactly, but grouped in thirteen 4-week periods, there was often some uncertainty as to the accident interval to which they belonged. But it was always possible to locate the longer periods of absence such as the 10—14 days annual holiday.

of small liability. The table compares the observed frequencies with the theoretical obtained by integrating (lx) between appropriate limits; also the mean z'', and the ratio of the difference between mean (observed) and mean (theory) to the standard error. In the case of $n=2$, the theoretical mean and standard deviation of z'' are infinite so that the latter comparison cannot be made, but as far as can be judged from a sample of 50, the correspondence in the frequencies appears very close. For $n=4$ and $n=6$ the suggestion of too few observed values of z'' in the first group would need confirmation from fuller data*; the mean values of z'' are

TABLE X.

Distribution of z''; Omnibuses and Pedestrians.

Intervals between Omnibuses						Intervals between Pedestrians		
$n=2$			$n=4$			$n=5$		
z''	Observation	Theory	z''	Observation	Theory	z''	Observation	Theory
0— ·5	4	9·3	0— ·1	3	7·0	0—·1	4	3·2
·5— 1·0	7·5	4·7	·1— ·2	1	4·8	·1—·2	3	2·0
1·0— 1·5	1·5	2·8	·2— ·3	2	3·5	·2—·3	—	1·3
1·5— 2·0	3	1·9	·3— ·4	5	2·5	·3—·4	—	0·9
2·0— 2·5	2	1·3	·4— ·5	2	1·9	·4—·5	1	0·6
2·5— 3·0	—	1·0	·5—1·0	6	4·8	·5—·6	1	0·4
3·0— 4·0	2	1·4	1·0—1·5	3	1·7	·6—·7	1	0·3
4·0— 5·0	4	0·9	1·5—2·0	2	0·8	Greater than ·7	—	1·3
5·0— 6·0	2	0·7	2·0—5·0	2	0·9	—	—	—
6·0— 7·0	—	0·5	5·0—8·0	1	0·1	—	—	—
7·0— 8·0	—	0·4	Greater than 8·0	1	—	—	—	—
8·0— 9·0	—	0·3	—	—	—	—	—	—
9·0—10·0	—	0·3	—	—	—	—	—	—
Greater than 10·0	2	2·5						
Total	28	28·0	Total	28	28·0	Total	10	10·0
Mean z''	—	—		1·258	·500		·210	·333
Difference in terms of standard error				+4·6			−·82	

however quite in agreement with theory as far as can be judged with such large standard errors.

To obtain some further idea of the sensitiveness of the test, we have applied it to the time intervals observed between the passing of consecutive omnibuses on a London bus route. These intervals are not of course random, the omnibuses

* In general there could be no closed interval, b, the times between accidents varying from 0 to over 180 days, but it is possible that after an accident the worker is for a day or two more cautious, and this reduces the chance of very small intervals and therefore small values of z''.

having started at fixed intervals (apparently of 3 minutes) from three-quarters of an hour to an hour before the observations were made. During this time traffic delays, etc. would have superposed chance positive and negative variations on to the scheduled intervals, so that in some cases these had been reduced to a few seconds and in others had been doubled. Table X gives the values of z'' calculated for 28 sets of $n=2$ and 28 of $n=4$. The fit for $n=2$ is far less satisfactory than in the case of the accident intervals, and appears even worse for $n=4$, where the mean z'' exceeds the theoretical value by 4·6 times its standard error. Finally in the same table are given 10 values of z'' calculated from sets of 5 intervals of time observed between pedestrians passing a fixed point on a roadway; here, as far as can be judged, agreement is satisfactory, suggesting that the intervals were random.

There has not been time before going to press to extend these results, but inadequate as the numbers are, they certainly suggest that the method may be of value in testing the randomness of intervals in many problems, and further that the test will be sensitive even for pairs of consecutive intervals ($n = 2$) provided that the number of such pairs available is large enough.

V. Conclusion.

The main problem that we have considered is that of determining whether it is probable that a given sample, taken as a whole, has been drawn from a specified population. From the idea of the representation of a sample of n as a point in multiple space, there follows at once the conception that the discriminating criteria may be regarded as associated with surfaces in this space which divide regions containing points for which the hypothesis to be tested as to the sample's origin is less probable from those for which it is more probable. Regarded from this point of view it is seen that in any given problem there need be no single system of surfaces which it is "best" to make use of. The system adopted will provide a numerical measure, and this must be coordinated in the mind of the statistician with a clear understanding of the process of reasoning on which the test is based. We have endeavoured to connect in a logical sequence several of the most simple tests, and in so doing have found it essential to make use of what R. A. Fisher has termed "the principle of likelihood." The process of reasoning, however, is necessarily an individual matter, and we do not claim that the method which has been most helpful to ourselves will be of greatest assistance to others. It would seem to be a case where each individual must reason out for himself his own philosophy.

The differences in method of approach have been illustrated in the two problems that have been described as the testing of Hypothesis A and Hypothesis B. Both problems suppose a knowledge of the form or shape of the population sampled, A supposes its standard deviation known and B does not. In both cases there appear to be four or five methods open to the statistician; if properly interpreted we should not describe one as more *accurate* than another, but according to the

problem in hand should recommend this one or that as providing information which is more *relevant* to the purpose. The principles involved have been illustrated in the case of samples from normal, rectangular and exponential populations, for each of which analogous tests exist. Most attention has been given to the normal case, simply because the β's of the populations of experience do tend to cluster round the Gaussian point. New tables have been computed for what is termed the P_A test of Hypothesis A, and illustrations of some of the various types of problems in which it could be employed have been given.

Now these tests, whether applicable to the normal, the rectangular or the exponential population, have all assumed both that there is perfect random sampling and that while there is uncertainty as to position and scale there is none as to the shape of the population curve. In following what would appear a more efficient system of reasoning, it must be remembered that the ideal situation is not quite that of practical reality, and we must be careful that what is gained on the swings is not being lost on the roundabouts. It would be useless to emphasise the gain in relevance of a $P_z = \cdot 08$ obtained from Student's Tables over a $P = \cdot 15$ obtained from Sheppard's Tables, if in fact the former were in error by an amount of the order of $\cdot 05$ owing to a faulty assumption that the population sampled was normal. Owing to the rapidity with which the distribution of means tends to normality, as n increases, the P obtained by entering Sheppard's Tables with $z = m/s$ has a definite meaning if properly interpreted, even if the population sampled be skew. But by using P_z we may be obtaining a false impression of the accuracy of the criterion upon which our judgment is to be based. In practice in dealing with small samples it is generally not possible to be certain of the form of the population distribution from internal evidence. But from other considerations, from previous experience in similar problems and so forth, we may have good grounds for believing that the distribution is roughly of a certain type, and if skew for knowing the direction in which it will be more sharply limited. There will undoubtedly be regions in the β_1, β_2 population field surrounding the normal, the rectangular, and the exponential points and any other points for which the problem can be completely solved, within which the criteria appropriate for these spots will be adequate. It is therefore a problem of first importance to ascertain what is the extent of these regions of adequacy, and failing a fresh advance in theory* the most straightforward method of doing this is to take experimental soundings in the β_1, β_2 field. In this process we have not at present gone very far, but have shown that

(*a*) At $\beta_1 = 0{\cdot}2$, $\beta_2 = 3{\cdot}2$, the normal theory for samples of 10 does provide criteria which for practical purposes are probably adequate in the case of the distribution of m, of s, of z and for the P_A contours.

* It is of course a serious limitation to the practical value of the method of likelihood, that it appears only to provide readily soluble equations in certain simple cases. It is not until the samples are large enough for the application of the (P, χ^2) group test that the problem can be solved in a more general manner.

(*b*) At $\beta_1 = 0$, $\beta_2 = 1{\cdot}8$, the z-distribution for samples of 4 and of 10 appears adequate, but the normal theory quite fails for s and P_λ. In this direction an intermediate sounding is required.

(*c*) Some experimental work at present in progress suggests that the normal theory will also be adequate for s and z for samples of 5 and 10 from a symmetrical leptokurtic curve with β_2 as great as 4·2.

(*d*) The results obtained in connection with the exponential population are quite out of touch with the symmetrical cases, but the distributions reached in that case will probably be adequate for a fairly wide range of skew populations with β_1 and β_2 in the region surrounding the point (4, 9).

The limitation implied by the assumption of perfect random sampling must not of course be overlooked; this is of particular importance in the case of very small samples. It has been shown that there may be some latitude as to the exact shape of the population curve associated with each hypothesis, but it has been assumed that in any event the sample is a random one. It is true that the causes underlying the bias or selection which occurs in the drawing of a sample from Π may be such that the sample can be regarded as randomly drawn from Π'; but if a biased sample from an approximately normal population can only be considered as a random sample from a J-shaped curve, it is clear that our methods of comparison will be inadequate. The bias may consist only of a few errors of observation or of crudeness of grouping, but the smaller the sample the more important will be the effect.

The answer to criticism from this source is that the tests should only be regarded as tools which must be used with discretion and understanding, and not as instruments which in themselves give the final verdict. It is impossible to escape from the difficulties; they make it desirable to avoid the use of very small samples whenever possible, but do not prevent some conclusions of value being drawn when fuller data are not available. If the criteria that have been discussed indicate that the supposed population of origin is improbable in comparison with some other alternative, that is a sign which is not to be disregarded. But we must not discard the original hypothesis until we have examined the alternative suggested, and have satisfied ourselves that it does involve a change in the real underlying factors in which we are interested; until we have assured ourselves that the alternative hypothesis is not error in observation, error in record, variation due to some outside factor that it was believed had been controlled, or to any one of many causes which when they chance to modify even one observation in a sample of five may be of serious consequence.

We must not conclude without thanking Miss Ida McLearn for her diagrams, Miss M. Page, without whose assistance in computing the completion of the paper in its present form must have been indefinitely delayed, and Mrs L. J. Comrie who carried out the final stage in the tabling of P_λ.

VI. APPENDIX: TABLES OF P_λ.

P_λ is given by the relation

$$P_\lambda = \int_\omega \frac{n^{\frac{n}{2}}}{\sqrt{\pi}\, 2^{\frac{n-2}{2}}\, \Gamma\left(\frac{n-1}{2}\right)} S^{n-2} e^{-\frac{n}{2}(M^2+S^2)}\, dM\, dS \quad \text{.....(xxi } bis\text{)},$$

where the integral is to be taken outside the curve in the (M, S)-plane upon which λ is constant, and

$$\lambda = S^n e^{-\frac{n}{2}(M^2+S^2-1)} \quad \text{.....(xix } bis\text{)}.$$

We may write

$$\log_{10}\lambda = \frac{n}{2}\{\log_{10} S^2 - (M^2+S^2-1)\log_{10} e\} = \frac{n}{2}\{\log_{10} e - k\},$$

and obtain for the equations of the contours

$$(M^2+S^2)\log_{10} e - \log_{10} S^2 = k \quad \text{.....(xx } bis\text{)}.$$

Values of M for equidistant values of S* were first computed for the curves corresponding to a number of values of k between ·45 and 2·40. Some of the contours have been drawn in Figures 13 and 14. The lowest possible value of k corresponding to the centre of the system, $M=0$, $S=1$, $\lambda=1$, is

$$k = \log_{10} e = \cdot 434{,}2945.$$

If S_1 and S_2 be the points at which a contour cuts the axis of S, if M_s indicates the value of M at the point on the curve corresponding to a given S, and if we write $x_s = \sqrt{n}\, M_s$, then the integral (xxi *bis*) can be put into the form

$$\tfrac{1}{2}(1-P_\lambda) = \int_{S_1}^{S_2} \frac{n^{\frac{n-1}{2}}}{2^{\frac{n-3}{2}}\, \Gamma\left(\frac{n-1}{2}\right)} S^{n-2} e^{-\frac{n}{2}S^2}\, (\tfrac{1}{2}\alpha_s)\, dS \quad \text{.....(lxiii)},$$

where

$$\tfrac{1}{2}(1+\alpha_s) = \int_{-\infty}^{x_s} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\, dx,$$

and is obtained by entering Sheppard's Tables with $x_s = \sqrt{n}\, M_s$. The values of the integral (lxiii) were then calculated by quadrature. The task of calculating P_λ for an adequate framework of values of k and for $n=3$ to $n=50$ would have been very great, but it was discovered that the ratios P_λ/λ for a given k are very nearly independent of n. Table XI gives the values of P_λ actually computed divided by the corresponding λ. It was estimated that the computed value of P_λ might be in error by several units in the fifth decimal place, and this limits the number of significant figures that can be given for the ratio as k increases. It also means that the final figures in the tabled ratios are not exact. The results however appeared to justify the following procedure:

(*a*) To obtain interpolated series of ratios for $n=3$, 4 and 10 from the ten or eleven values computed in each case.

* The argument interval for S varied from ·005 to ·04.

TABLE XI.

Computed Values of P_λ/λ.

k	$n=3$	$n=4$	$n=5$	$n=10$	$n=20$	$n=50$
log e	1·0000	1·0000	1·0000	1·0000	1·0000	1·0000
·45	—	1·0169	—	1·0170	—	1·0170
·50	1·0743	—	1·0736	1·0731	1·0732	1·072
·55	—	—	—	1·1327	—	1·135⁻
·60	1·1991	1·1977	—	1·1959	1·195⁻	—
·65	—	—	—	1·2622	—	—
·70	—	1·3363	—	1·332	1·33	—
·75	1·4151	1·4117	—	—	—	—
·80	1·4957	1·4921	1·4899	1·485⁺	—	—
·90	1·6721	1·6672	—	1·66	—	—
1·00	—	1·864	—	1·85	—	—
1·20	2·3432	2·335⁻	2·330	—	—	—
1·50	3·292	3·286	—	—	—	—
1·80	4·638	4·62	—	—	—	—
2·40	9·42	—	—	—	—	—
3·00	18·53	—	—	—	—	—

(*b*) To extrapolate for high values of k for $n = 10$ and 4, from $n = 4$ and 3 respectively, in view of the fact that as P_λ decreases only a few significant figures are needed in the ratio.

(*c*) With the help of these "frames" and the isolated values for $n = 5$, to interpolate appropriate ratios for $n = 5$, 6, 7, 8 and 9.

(*d*) Since the ratios at $n = 10$ give accurately the three values of P_λ computed for $n = 20$ and the three for $n = 50$, to assume that the ratios for $n = 10$ will be adequate for all values of n in the range 10 to 50.

To obtain P_λ it was only necessary to compute $\log_{10}\lambda = \frac{n}{2}(\log_{10} e - k)$, add to it the logarithm of the ratio P_λ/λ and find the anti-logarithm. The P_λ tables given below were calculated in this way. They are given to four decimal places, but the final figure may often be in error by a single unit, although not, it is believed, by as much as two units. After the work was commenced it was realised that the variable k was not perhaps the best to have taken from the point of view of interpolation in the tables; k^2 would probably have been better. It has been necessary to change the interval from ·01 to ·05 at $k = ·65$, and from ·05 to ·30 at $k = 1·50$. For high values of n and low values of k the tables are not easy to use if four-figure accuracy is required, but for most practical purposes it is sufficient to obtain P_λ to two decimal places only, and in this there is no difficulty.

Method of entry.

Having obtained from the data of the problem under consideration $M = m/\sigma$ and $S = s/\sigma$, k may either be calculated from (xx *bis*), or generally obtained with sufficient accuracy from one or other of the Figures 13 and 14. P_λ is then found by entering the tables with n and k. Examples of the use of the tables have been given in Section (9) above.

Extension of the Tables.

On the assumption that the ratios P_λ/λ for $n = 10$ will continue to hold good for any value of n greater than 50, we may proceed as follows:

(*a*) Find the appropriate k, (*b*) calculate λ from the relation

$$\log_{10}\lambda = \frac{n}{2}(\log_{10} e - k),$$

(*c*) multiply the ratio P_λ/λ given in Table XII below by λ, and so obtain P_λ.

TABLE XII.

Values of P_λ/λ for $n = 10$ and beyond.

k	P_λ/λ	k	P_λ/λ	k	P_λ/λ	k	P_λ/λ
·435	1·0008	·500	1·0731	·565	1·1514	·630	1·2353
·440	1·0062	·505	1·0789	·570	1·1576	·635	1·2420
·445	1·0116	·510	1·0847	·575	1·1639	·640	1·2488
·450	1·0170	·515	1·0906	·580	1·1702	·645	1·2555
·455	1·0224	·520	1·0965	·585	1·1766	·650	1·2623
·460	1·0279	·525	1·1024	·590	1·1830	·70	1·332
·465	1·0334	·530	1·1084	·595	1·1894	·75	1·407
·470	1·0390	·535	1·1144	·600	1·1959	·80	1·485+
·475	1·0446	·540	1·1205	·605	1·2024	·85	1·569
·480	1·0502	·545	1·1266	·610	1·2089	·90	1·658
·485	1·0559	·550	1·1328	·615	1·2155	·95	1·753
·490	1·0616	·555	1·1390	·620	1·2221	1·00	1·855
·495	1·0673	·560	1·1452	·625	1·2287	—	—

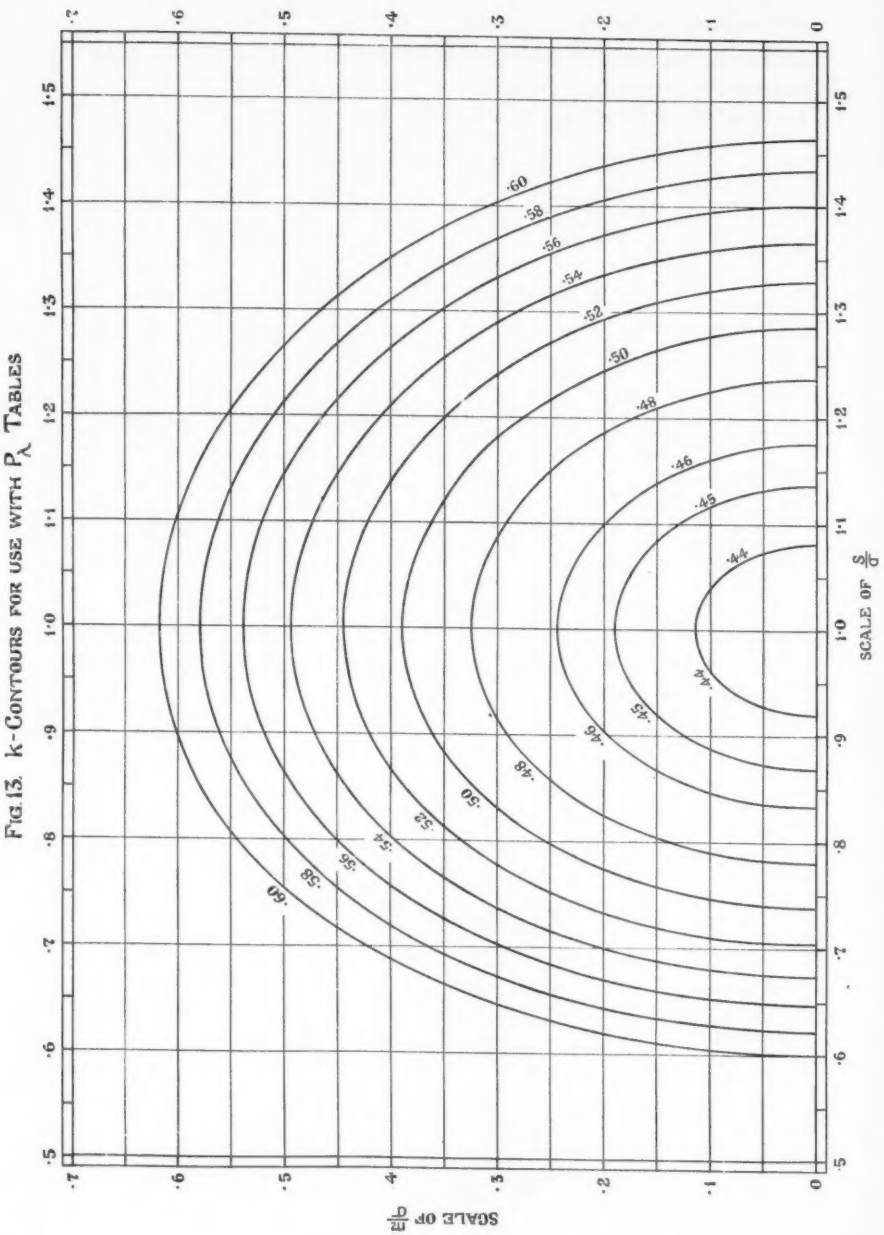


FIG 13. k-CONTOURS FOR USE WITH P_λ TABLES

FIG 14. k-CONTOURS FOR USE WITH P_λ TABLES

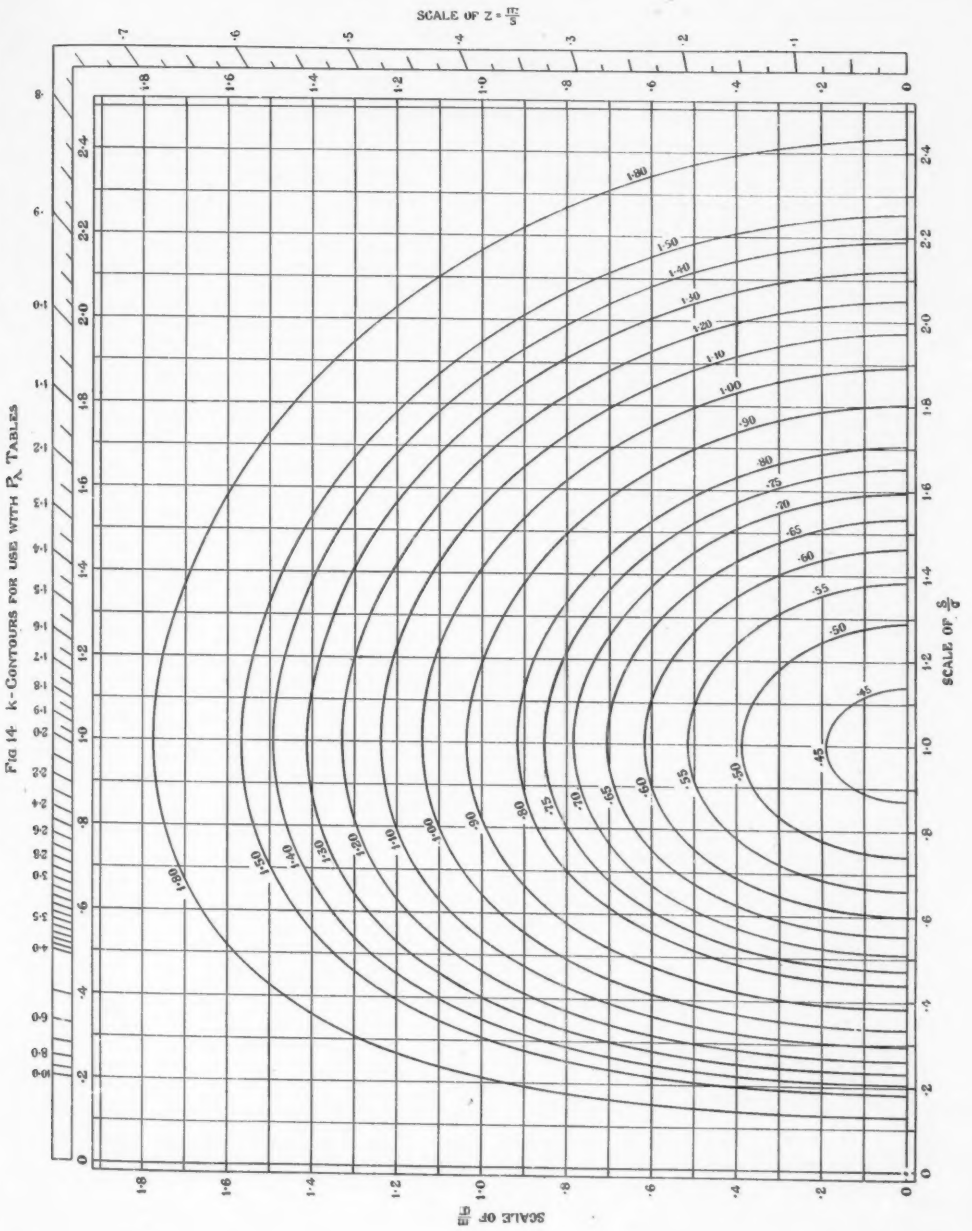


FIG 14. k-CONTOURS FOR USE WITH P_A TABLES

TABLES OF P_A.

Size of Sample, n.

k	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*k*
·44	·9865	·9800	·9736	·9672	·9609	·9546	·9484	·9421	·9360	·9298	·9238	·9177	·44
·45	·9634	·9460	·9291	·9124	·8961	·8800	·8642	·8487	·8335	·8186	·8039	·7895	·45
·46	·9408	·9132	·8866	·8607	·8356	·8113	·7876	·7646	·7423	·7207	·6996	·6792	·46
·47	·9189	·8817	·8461	·8120	·7792	·7478	·7177	·6888	·6611	·6344	·6089	·5844	·47
·48	·8975	·8513	·8075	·7660	·7267	·6894	·6541	·6205	·5887	·5585	·5299	·5027	·48
·49	·8766	·8219	·7707	·7227	·6778	·6356	·5961	·5590	·5243	·4917	·4612	·4325	·49
·50	·8562	·7935	·7355	·6818	·6321	·5859	·5432	·5036	·4669	·4329	·4014	·3721	·50
·51	·8363	·7661	·7019	·6432	·5895	·5402	·4951	·4537	·4159	·3811	·3493	·3202	·51
·52	·8168	·7396	·6699	·6068	·5497	·4980	·4512	·4088	·3704	·3356	·3040	·2755	·52
·53	·7978	·7140	·6393	·5724	·5126	·4591	·4112	·3683	·3299	·2954	·2646	·2370	·53
·54	·7792	·6894	·6101	·5401	·4781	·4233	·3748	·3318	·2938	·2601	·2303	·2039	·54
·55	·7611	·6656	·5823	·5096	·4459	·3903	·3416	·2990	·2617	·2290	·2005	·1755	·55
·56	·7434	·6426	·5558	·4808	·4159	·3598	·3113	·2694	·2331	·2017	·1745	·1510	·56
·57	·7261	·6204	·5304	·4536	·3879	·3318	·2837	·2427	·2076	·1776	·1519	·1299	·57
·58	·7092	·5990	·5063	·4280	·3618	·3059	·2586	·2187	·1849	·1563	·1322	·1118	·58
·59	·6927	·5784	·4832	·4038	·3374	·2820	·2357	·1970	·1647	·1376	·1151	·0962	·59
·60	·6766	·5584	·4611	·3810	·3147	·2600	·2148	·1775	·1467	·1212	·1001	·0827	·60
·61	·6608	·5391	·4401	·3594	·2935	·2397	·1958	·1599	·1306	·1067	·0872	·0712	·61
·62	·6454	·5205	·4201	·3391	·2737	·2210	·1784	·1441	·1163	·0939	·0759	·0613	·62
·63	·6305	·5026	·4009	·3199	·2553	·2037	·1626	·1298	·1036	·0827	·0660	·0527	·63
·64	·6158	·4852	·3826	·3018	·2381	·1878	·1482	·1169	·0923	·0728	·0575	·0453	·64
·65	·6015	·4685	·3652	·2847	·2220	·1732	·1350	·1053	·0822	·0641	·0500	·0390	·65
·70	·5348	·3931	·2892	·2129	·1567	·1153	·0849	·0625	·0461	·0339	·0250	·0184	·70
·75	·4756	·3299	·2291	·1591	·1106	·0768	·0534	·0371	·0258	·0180	·0125	·0087	·75
·80	·4229	·2769	·1815	·1190	·0781	·0512	·0336	·0221	·0145	·0095	·0062	·0041	·80
·85	·3762	·2325	·1439	·0891	·0551	·0341	·0211	·0131	·0081	·0050	·0031	·0019	·85
·90	·3347	·1952	·1140	·0666	·0389	·0228	·0133	·0078	·0046	·0027	·0016	·0009	·90
·95	·2979	·1640	·0904	·0499	·0275	·0152	·0084	·0046	·0026	·0014	·0008	·0004	·95
1·00	·2651	·1377	·0717	·0373	·0194	·0101	·0053	·0027	·0014	·0007	·0004	·0002	1·00
1·05	·2359	·1157	·0568	·0279	·0137	·0068	·0033	·0016	·0008	·0004	·0002	·0001	1·05
1·10	·2100	·0972	·0450	·0209	·0097	·0045	·0021	·0010	·0005	·0002	·0001	·0000	1·10
1·15	·1869	·0817	·0357	·0156	·0069	·0030	·0013	·0006	·0003	·0001	·0000	—	1·15
1·20	·1664	·0687	·0284	·0117	·0048	·0020	·0008	·0003	·0001	·0001	—	—	1·20
1·25	·1482	·0578	·0226	·0088	·0034	·0013	·0005	·0002	·0001	·0000	—	—	1·25
1·30	·1319	·0486	·0179	·0066	·0024	·0009	·0003	·0001	·0000	—	—	—	1·30
1·35	·1174	·0409	·0142	·0049	·0017	·0006	·0002	·0001	—	—	—	—	1·35
1·40	·1046	·0344	·0113	·0037	·0012	·0004	·0001	·0000	—	—	—	—	1·40
1·45	·0932	·0289	·0090	·0028	·0009	·0003	·0001	—	—	—	—	—	1·45
1·50	·0830	·0243	·0071	·0021	·0006	·0002	·0001	—	—	—	—	—	1·50
1·80	·0415	·0086	·0018	·0004	·0001	·0000	·0000	—	—	—	—	—	1·80
2·10	·0206	·0030	·0004	·0001	·0000	—	—	—	—	—	—	—	2·10
2·40	·0106	·0011	·0001	·0000	—	—	—	—	—	—	—	—	2·40
2·70	·0053	·0004	·0000	—	—	—	—	—	—	—	—	—	2·70
3·00	·0026	·0001	—	—	—	—	—	—	—	—	—	—	3·00

Size of Sample, n.

k	15	16	17	18	19	20	21	22	23	24	25	26	k
·44	·9117	·9057	·8998	·8939	·8881	·8822	·8765	·8707	·8650	·8593	·8537	·8481	·44
·45	·7754	·7615	·7478	·7344	·7213	·7084	·6956	·6832	·6709	·6589	·6471	·6355	·45
·46	·6594	·6402	·6215	·6034	·5858	·5687	·5521	·5360	·5204	·5052	·4905	·4762	·46
·47	·5608	·5382	·5166	·4958	·4758	·4566	·4382	·4206	·4037	·3874	·3718	·3568	·47
·48	·4770	·4525	·4293	·4073	·3864	·3666	·3478	·3300	·3131	·2970	·2818	·2674	·48
·49	·4057	·3805	·3568	·3347	·3139	·2944	·2761	·2589	·2429	·2278	·2136	·2004	·49
·50	·3450	·3199	·2966	·2750	·2549	·2364	·2191	·2032	·1884	·1746	·1619	·1501	·50
·51	·2934	·2690	·2465	·2259	·2071	·1898	·1739	·1594	·1461	·1339	·1227	·1125	·51
·52	·2496	·2261	·2049	·1856	·1682	·1524	·1381	·1251	·1133	·1027	·0930	·0843	·52
·53	·2123	·1901	·1703	·1525	·1366	·1224	·1096	·0982	·0879	·0787	·0705	·0632	·53
·54	·1806	·1599	·1416	·1253	·1110	·0983	·0870	·0770	·0682	·0604	·0535	·0473	·54
·55	·1536	·1344	·1177	·1030	·0901	·0789	·0691	·0604	·0529	·0463	·0405	·0355	·55
·56	·1306	·1130	·0978	·0846	·0732	·0634	·0548	·0474	·0410	·0355	·0307	·0265	·56
·57	·1111	·0950	·0813	·0695	·0595	·0509	·0435	·0372	·0318	·0272	·0233	·0199	·57
·58	·0945	·0799	·0676	·0571	·0483	·0409	·0345	·0292	·0247	·0209	·0177	·0149	·58
·59	·0804	·0672	·0562	·0470	·0392	·0328	·0274	·0229	·0192	·0160	·0134	·0112	·59
·60	·0684	·0565	·0467	·0386	·0319	·0263	·0218	·0180	·0149	·0123	·0101	·0084	·60
·61	·0582	·0475	·0388	·0317	·0259	·0212	·0173	·0141	·0115	·0094	·0077	·0063	·61
·62	·0495	·0399	·0323	·0260	·0210	·0170	·0137	·0111	·0089	·0072	·0058	·0047	·62
·63	·0421	·0336	·0268	·0214	·0171	·0136	·0109	·0087	·0069	·0055	·0044	·0035	·63
·64	·0358	·0282	·0223	·0176	·0139	·0110	·0086	·0068	·0054	·0042	·0034	·0026	·64
·65	·0304	·0237	·0186	·0145	·0113	·0088	·0069	·0054	·0042	·0033	·0025	·0020	·65
·70	·0135	·0100	·0073	·0054	·0040	·0029	·0022	·0016	·0012	·0009	·0006	·0005	·70
·75	·0060	·0042	·0029	·0020	·0014	·0010	·0007	·0005	·0003	·0002	·0002	·0001	·75
·80	·0027	·0018	·0012	·0008	·0005	·0003	·0002	·0001	·0001	·0001	·0000	·0000	·80
·85	·0012	·0007	·0005	·0003	·0002	·0001	·0001	·0000	·0000	·0000	—	—	·85
·90	·0005	·0003	·0002	·0001	·0001	·0000	·0000	—	—	—	—	—	·90
·95	·0002	·0001	·0001	·0000	·0000	—	—	—	—	—	—	—	·95
1·00	·0001	·0001	·0000	—	—	—	—	—	—	—	—	—	1·00
1·05	·0000	·0000	—	—	—	—	—	—	—	—	—	—	1·05

Size of Sample, n.

k	27	28	29	30	31	32	33	34	35	36	37	38	k
·44	·8426	·8371	·8316	·8262	·8207	·8154	·8100	·8047	·7994	·7942	·7890	·7839	·44
·45	·6241	·6129	·6020	·5912	·5806	·5702	·5600	·5499	·5401	·5304	·5209	·5116	·45
·46	·4623	·4488	·4357	·4230	·4107	·3987	·3871	·3758	·3648	·3542	·3439	·3339	·46
·47	·3425	·3287	·3154	·3027	·2905	·2788	·2676	·2568	·2465	·2365	·2270	·2179	·47
·48	·2537	·2407	·2283	·2166	·2055	·1950	·1850	·1755	·1665	·1580	·1499	·1422	·48
·49	·1879	·1762	·1653	·1550	·1454	·1364	·1279	·1199	·1125	·1055	·0989	·0928	·49
·50	·1392	·1290	·1196	·1109	·1028	·0954	·0884	·0820	·0760	·0705	·0653	·0606	·50
·51	·1031	·0945	·0866	·0794	·0728	·0667	·0611	·0560	·0513	·0471	·0431	·0395	·51
·52	·0764	·0692	·0627	·0568	·0515	·0466	·0423	·0383	·0347	·0314	·0285	·0258	·52
·53	·0566	·0507	·0454	·0407	·0364	·0326	·0292	·0262	·0234	·0210	·0188	·0168	·53
·54	·0419	·0371	·0329	·0291	·0258	·0228	·0202	·0179	·0158	·0140	·0125	·0110	·54
·55	·0311	·0272	·0238	·0208	·0182	·0160	·0140	·0122	·0107	·0094	·0082	·0072	·55
·56	·0230	·0199	·0172	·0149	·0129	·0112	·0097	·0084	·0072	·0063	·0054	·0047	·56
·57	·0170	·0146	·0125	·0107	·0091	·0078	·0067	·0057	·0049	·0042	·0036	·0031	·57
·58	·0126	·0107	·0090	·0076	·0065	·0055	·0046	·0039	·0033	·0028	·0024	·0020	·58
·59	·0094	·0078	·0065	·0055	·0046	·0038	·0032	·0027	·0022	·0019	·0016	·0013	·59
·60	·0069	·0057	·0047	·0039	·0032	·0027	·0022	·0018	·0015	·0012	·0010	·0008	·60
·61	·0051	·0042	·0034	·0028	·0023	·0019	·0015	·0012	·0010	·0008	·0007	·0006	·61
·62	·0038	·0031	·0025	·0020	·0016	·0013	·0011	·0009	·0007	·0006	·0004	·0004	·62
·63	·0028	·0022	·0018	·0014	·0011	·0009	·0007	·0006	·0005	·0004	·0003	·0002	·63
·64	·0021	·0016	·0013	·0010	·0008	·0006	·0005	·0004	·0003	·0002	·0002	·0002	·64
·65	·0015	·0012	·0009	·0007	·0006	·0004	·0003	·0003	·0002	·0002	·0001	·0001	·65
·70	·0003	·0003	·0002	·0001	·0001	·0001	·0001	·0000	·0000	·0000	·0000	·0000	·70
·75	·0001	·0001	·0000	·0000	·0000	·0000	·0000	—	—	—	—	—	·75
·80	·0000	·0000	—	—	—	—	—	—	—	—	—	—	·80

Size of Sample, n.

k	*39*	*40*	*41*	*42*	*43*	*44*	*45*	*46*	*47*	*48*	*49*	*50*	*k*
·*44*	·7787	·7736	·7686	·7635	·7585	·7536	·7486	·7437	·7389	·7340	·7292	·7244	·*44*
·*45*	·5024	·4934	·4846	·4759	·4673	·4590	·4507	·4427	·4347	·4269	·4193	·4118	·*45*
·*46*	·3241	·3147	·3055	·2966	·2879	·2795	·2714	·2635	·2558	·2483	·2411	·2341	·*46*
·*47*	·2091	·2007	·1926	·1848	·1774	·1703	·1634	·1568	·1505	·1444	·1386	·1330	·*47*
·*48*	·1349	·1280	·1214	·1152	·1093	·1037	·0984	·0933	·0886	·0840	·0797	·0756	·*48*
·*49*	·0870	·0816	·0766	·0718	·0673	·0632	·0592	·0556	·0521	·0489	·0458	·0430	·*49*
·*50*	·0562	·0521	·0483	·0448	·0415	·0385	·0357	·0331	·0307	·0284	·0264	·0244	·*50*
·*51*	·0362	·0332	·0304	·0279	·0256	·0234	·0215	·0197	·0180	·0165	·0152	·0139	·*51*
·*52*	·0234	·0212	·0192	·0174	·0158	·0143	·0129	·0117	·0106	·0096	·0087	·0079	·*52*
·*53*	·0151	·0135	·0121	·0108	·0097	·0087	·0078	·0070	·0062	·0056	·0050	·0045	·*53*
·*54*	·0097	·0086	·0076	·0068	·0060	·0053	·0047	·0042	·0037	·0033	·0029	·0026	·*54*
·*55*	·0063	·0055	·0048	·0042	·0037	·0032	·0028	·0025	·0022	·0019	·0017	·0015	·*55*
·*56*	·0041	·0035	·0030	·0026	·0023	·0020	·0017	·0015	·0013	·0011	·0010	·0008	·*56*
·*57*	·0026	·0022	·0019	·0016	·0014	·0012	·0010	·0009	·0007	·0006	·0005	·0005	·*57*
·*58*	·0017	·0014	·0012	·0010	·0009	·0007	·0006	·0005	·0004	·0004	·0003	·0003	·*58*
·*59*	·0011	·0009	·0008	·0006	·0005	·0004	·0004	·0003	·0003	·0002	·0002	·0002	·*59*
·*60*	·0007	·0006	·0005	·0004	·0003	·0003	·0002	·0002	·0002	·0001	·0001	·0001	·*60*
·*61*	·0005	·0004	·0003	·0002	·0002	·0002	·0001	·0001	·0001	·0001	·0001	·0000	·*61*
·*62*	·0003	·0002	·0002	·0002	·0001	·0001	·0001	·0001	·0001	·0000	·0000	—	·*62*
·*63*	·0002	·0002	·0001	·0001	·0001	·0001	·0000	·0000	·0000	—	—	—	·*63*
·*64*	·0001	·0001	·0001	·0001	·0000	·0000	—	—	—	—	—	—	·*64*
·*65*	·0001	·0001	·0000	·0000	—	—	—	—	—	—	—	—	·*65*

N.B. I feel it necessary to make a brief comment on the authorship of this paper. Its origin was a matter of close co-operation, both personal and by letter, and the ground covered included the general ideas and the illustration of these by sampling from a normal population. A part of the results reached in common are included in Chapters I, II and V. Later I was much occupied with other work, and therefore unable to co-operate. The experimental work, the calculation of tables and the developments of the theory of Chapters III and IV are due solely to Dr Egon S. Pearson.

J. NEYMAN.